

スキーマの形成

7月号の古山先生の「音源の分離」は多くの読者の目を惹いたことと思います。「ピッチトラッキング」「テンポトラッキング」と並んで「スペクトラムトラッキング」は夢のハイテクとして未だに完成していません。国内でもかつて大阪大学の片寄研究室の研究が先駆的でした。所謂「ロボット聴覚」という「音楽認知モデル」の研究の難しさは、言語認識の研究の数百倍になると言われています。

今日のパソコンには音声認識ソフトがほぼ標準で装備されています。或いは逆に読み上げソフトもかなりのレベルになっていますが、これらのソフトの基幹をなす研究は「パターン認識」と「ツリー構造スキーマ」なのです。音声認識ソフトの場合最初にユーザーの声やクセを学習させて精度を高めます。この原理に基づいた音楽認知ソフトも1990年代にはICMC(国際コンピュータミュージック学会)で多く発表されました。パターン認識は出来るだけ短い時間内にサンプルと対象の共通部分の多いモノを選ぶという謂わばサンプル数と演算速度のレベルに依存します。これに予測能力を持たせて先回りしてサンプルを用意するものまで現れました。原田技研の原田豊成氏の発明による「音感と歌唱の上達器」ではたった2つのICチップで2セントの精度でアナログ音声のピッチが検出できます。それでも、複数の声やパートを認知することは夢の夢です。本当に人間の認知能力というのは凄いと感心しますが、生まれてきたその日からその能力があったわけではなく、言語の習得と同じように単純なものから複雑なものへと次第に学習されて獲得される能力なのです。従って人工音楽認知ソフトにもワープロソフトのような学習機能が基本的な原理に包含されます。

この原理の裏には実際の音楽能力の形成に関わる音楽教育がヒントとして隠れています。残念なことに多くの音楽認知ソフトやハードは実際に音楽教育に携わった経験や、実際に演奏や創作活動に従事したことの無い研究者によって行われてきました。先に紹介した原田氏も医療機器界での発明家なので、CTの開発の先駆者ではありますが、音楽や音楽教育とは無縁でした。物理的現象としての「音」と心理的現象としての「音楽」は、「音声」と「言語」

の関係に似ています。言語学者が多く関わった漢字変換ソフトは誤変換の少ない使いやすいソフトですが、最近の変換ソフトは「予測変換」が常識です。この予測変換は文脈の前後関係から予測して適切な漢字の優先順位を決めるものですが、これが無いと迷文珍文が生まれるわけです。

この予測の精度はスキーマ(認知のための知識構造)で決まります。音楽にも「文脈」はあります。その文脈を正しく解釈して、予測を立てるという行為が「音楽を聴く」という行為なんです。この行為が成立するためには、「音楽的イディオム(慣用句)」を知識として形成しておく必要があります。「わからない音楽」はこの慣用句が通用しないケースで、いわば外国語を「音」としては聞けるが、「言語」としては理解できないのに似ています。外国語を理解するためには「音」の組み合わせに意味を持たせるスキーマが必要で、そのスキーマを形成する学習が幼児期に行われることが望ましいのです。

例えば幼児期に簡単な「アー」「ウー」から「アイウエオ・・・」の50音を日本語のスキーマとして形成した我々LとRを区別する発音や、ローマ字読みできないAとOの中間のような発音のスキーマを持ちませんので、英語やロシア語の発音を正しく聞き取るためのスキーマを改めて学習するのに大変苦労します。イタリア語は母音がローマ字と非常に近いので限りなく日本語の50音に近い近親感がありますし、ハワイのマウイ語は36音(5つの母音と7つの子音)しかありませんのでもっと我々には「聞き取りやすい」幼児語に近い言語です。

本稿バックナンバー25~29「データと情報」というテーマで述べましたように「音」は単独で情報を持つ場合もありますが、「音楽」は「音」をデータとした情報なのです。その情報は題名や調性、テンポ、演奏形態、歌詞、楽器などによって構成されますが、基本的には「音」の組み合わせで構成される「情報」です。

音楽教育の目的はこの「情報」解読のためのスキーマを形成させることであり、いかに多くの有効な「音楽的イディオム」を照合用テーブルに蓄えるかという質と量の教育でもあるのです。