

聞こえと無意識の科学

物理情報システム専攻 柏野 牧夫 研究室

柏野 牧夫 連携教授 1964年岡山県生まれ。東京大学大学院人文科学研究科修士課程修了。博士（心理学）。現在、NTTコミュニケーション科学基礎研究所上席特別研究員・人間情報研究部長、東京工業大学大学院総合理工学研究科物理情報システム専攻連携教授。



先生は人間の脳の無意識的な情報処理のメカニズムを明らかにしようとしている。特に聴覚について、錯覚を利用することでどのようにして主観的な聞こえというものが形成されるのか追究している。また先生は音楽を聴いたときの心地よさや魅力の形成についても、数理科学・神経生理学的なアプローチからモデルを立てて解析を進めている。

脳は無意識的な処理

仕掛けがわかっているにもかかわらず、無意識にそう見えてしまう。これは錯覚の醍醐味である。テレビ番組や書籍などで錯覚を体験したことがある人も多いだろう。私たちの脳を巧妙にだます錯覚だが、そもそも錯覚とはどのような現象なのだろうか。

錯覚とは、感覚器に異常がないにもかかわらず実際とは異なる知覚を得てしまう現象を指す。例えば、欠けたチーズ片のような黒い図形と、灰色の水玉のような図形がある（図1）。黒い図形は、単体では何を意味しているのかよくわからない。しかし、灰色の図形と重ねるとABCDと読むことができる。本来読めないはずのABCDの文字がこのように読める現象は錯覚の一例だと言える¹⁾。

脳は視覚情報の処理を行うことで、図形を知覚する。この過程は自動的かつ無意識的なものだ。

情報処理した結果が視覚に入ったものからずれると、脳は錯覚をおこすのである。

錯覚以外にも、脳が無意識的な処理を行なっているものがある。例えば音楽を聴いているときがそうだ。オーケストラの演奏を聴いて鳥肌が立ったり、赤ちゃんが子守唄を聴いて眠りに落ちてしまったりすることがある。こういった光景は日常的によく見られるが、このような心身の反応は意識しなくとも感じられる。これは、実際に音楽を聴いているときにも錯覚のときと同じく、脳が無意識的な処理を行なっているからである。

脳は無意識的な処理といっても、処理を行う脳の部位や処理の段階、連鎖的な反応は非常に多岐にわたっている。これらの処理に関するメカニズムには未解明な部分が多い。柏野先生はこうした脳の無意識的な処理のメカニズムを、主に聴覚の側面から解明しようと考えている。

1) <http://www.brl.ntt.co.jp/IllusionForum/v/findLetter1/ja/index.html>（※再生にはAdobe Flash Player環境を必要とする）

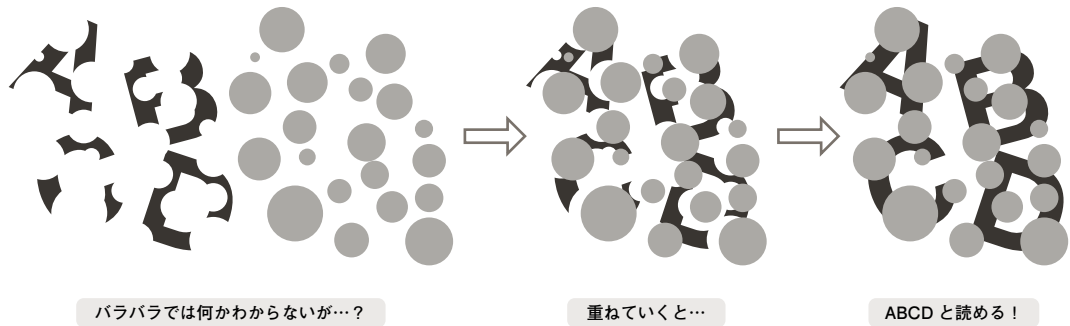


図1 錯視の一例

左側の図では黒色の図形と灰色の図形が分かれており、このままでは何を表しているかわからない。中心の図のように徐々に二つの図形を重ねていき、右側の図のように完全に重なると、初めてABCDと読むことができる。

錯覚と視覚的特性

ABCDの錯覚の例で、どのような原因で錯覚が起こるのか考えてみよう。重ねる前の図と重ねた後の図はどこが違うのだろうか。チーズのような黒い欠片と灰色の水玉、どちらも構成しているものは変わらない。しかし図形の位置関係は異なる。特に重ねた後の図は、チーズ片と水玉の図形がそれぞれの縁で接している。

縁というものは一般に物体の境界線を指す。今眺めている本誌や動いている人間といったどんなものでも、視覚的にその縁を認識することができる。縁を認識することによって、初めて背景から物体を切り取って区別できるのだ。

ここで縁が一定の長さで接している二つの物体が目前にあるとしよう。このとき私たちには一方が他方を部分的に隠しているように見える。例えば、両手の人差し指をクロスさせて見てみる。このとき誰も後ろにある方の指が二つに分断されているとは思わないだろう。後ろの物体が部分的に隠されているために、一つの物体として繋がっていると無意識的に判断するのである。

縁は見る人に対して形状を補完する余地を与える。これに加えて見え方の経験則などから、脳は水玉に隠された本来の形状を補完する。結果アルファベットが浮かび上がって見えるのである。

錯覚の例としてほかにも、「婦人と老婆」というだまし絵がある(図2)。後ろを向いた婦人、もしくは左向きの老婆として見るができる。

この絵には見え方が二つあるが、重ね合わせるように婦人と老婆を同時に見ることは困難だ。婦人であれば婦人として、老婆であれば老婆として見るができない。ただし、見え方を次々に変えることは可能である。こういった複数の解釈をもつ図形のことを多義図形と呼ぶ。また、このような見え方の解釈の選択性は一つの視覚的特性と言える。

さて、ここまで取り上げてきた錯覚、これらはすべて視覚についてのものであった。専門的にはこれらの錯覚を錯視と呼ぶ。実は錯覚というものは錯味や錯嗅というように、人間の五感それぞれについて存在している。その中の一つが聴覚における錯覚、錯聴である。

これから登場する錯聴現象は、脳がどのようにして音の解析や分析を行なっているのか、先生が実際に研究を進めているテーマである。



図2 婦人と老婆

19世紀からある多義図形の一例である。

錯聴とメカニズム

錯聴に関する脳のメカニズムを明らかにする前に、そもそも聴覚情報はどのように神経や脳で処理されるのかを説明する。

耳に入った音声はまず鼓膜を震わせる。鼓膜の振動は内耳の蝸牛に伝わり、そこで音声の周波数の分析がなされる。電気信号に置き換えられた後、神経細胞によって脳幹、内側膝状体のある視床を通り、最終的に聴覚野に届く（図3）。

聴覚情報は神経系の上記の経路を進んでいくにつれて、より細かく処理される。一般に神経系の経路を進んでいけばいくほど、前の段階で処理された情報を調整する側の部位となる。つまり聴覚野は視床の機能を調整し、視床は脳幹の機能を調整する役割を果たしているのだ。

■ 連続聴効果

駅のホームを電車が通過するとき、電車がホームに近づくにつれて音が徐々に大きくなり、遠ざかるにつれて徐々に小さくなる。このような音の特徴として連続性が挙げられる。音の連続性についての聴覚的特性の一つに、連続聴効果がある。

錯聴現象は具体的に紙面に記載できないため

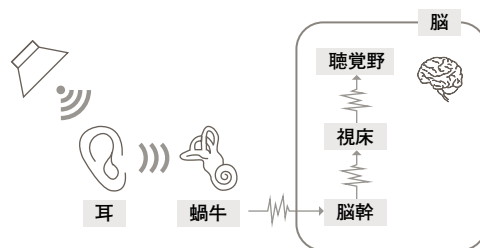


図3 聴覚情報処理のしくみ

ウェブページを利用する²⁾。URLのリンク先は連続聴効果についてのページである。リンク先のサイト、Illusion Forumには先生が制作した錯聴のデモンストレーションが多数公開されている。Illusion Forumでは、日常では体験しにくい錯聴現象を実際に体験することができる。

連続聴効果のページにある再生ボタンCの音では「エリーゼのために」というピアノ楽曲が流れるが、音量がゼロになっているところが周期的にあり、途切れ途切れで聴き取りにくい。一方、再生ボタンDの音では同じく「エリーゼのために」が流れるが、周期的にノイズが入っている。今度は、Cの音に比べ全体を通してなめらかに聴くことができる。両者の違いは楽曲と無関係なノイズがあるかどうかだけである（図4）。よって「エリーゼのために」が再生されている部分の音の情報はまったく変わらない。

滑らかに聴こえる理由は、ノイズがいわばABCDの錯視の例における水玉と同じ役目を果たすためである。錯視の例と同様、空白をほかのもので補完することでなめらかに聴くことができるのだ。

錯聴と錯視は時間的な現象、空間的な現象という点で異なっている。また、音と光の性質の違いもある。つまり錯聴の方はノイズ部分においてノイズと補完された音が同時に聴こえているが、錯視の方はABCDの文字の補完された部分が透過して見えているわけではない。けれども連続性という点で両者にはきれいな対応があることがわかる。

先ほど多義図形の錯視の例を紹介した。以下は、これに対応づけられる錯聴の例である選択的聴取と音脈分凝である。

図4 連続聴効果の例

ある楽曲の一部を定期的に無音にしたものとノイズを挿入したものを聴き比べることで、連続聴効果を実感することができる。

2) <http://www.brl.ntt.co.jp/IllusionForum/a/continuityIllusion/ja/index.html>（※再生にはAdobe Flash Player環境を必要とする）

■ 選択的聴取

同時に10人の話を聞き分けることができたという聖徳太子の伝説を知っている読者は多いかと思う。一般的に多人数の声を同時に聞くのは難しい。しかしながら話し声の絶えない人混みの中で、隣の友人の声を聞き、会話するようなことは、不自由なく行うことができる。ここでの聴覚の特性を選択的聴取という。

この現象は以下の簡単な実験として再現できる。多数の人が同時に別々の内容を話している音声を再生する。その中に数字を読み上げる女性がいる。被験者にはその人の声に注意を向けて音声を聴いてもらう。音声を聴いてもらった後に、音声で話していた合計の人数を尋ねるというものだ。

実験では被験者は読み上げられた数字の声を聴き取ることはできた。だが話者の人数については多くの場合、実際の人数が5人でも10人でも3、4人と答えた。

被験者は話者が2人までならその人数を正確に当てることができる。それは自分が聴いている声の話者に対して、注意を向けていない方の話者が排他的に定まるからだ。ところが3人以上の場合、注意を向けていない話者をそれぞれ特定することができない。すると被験者が答える人数と実際の人数とが合わなくなってくる。5人以上になるとほとんどの被験者は注意を向けていない話者の声をさばくことができない。だから3、4人と答えてしまうのだ。

この現象における脳の処理のメカニズムについて、ほかの研究室で行われた実験がある。聴覚野を含む側頭葉に数十本の電極を挿したうえで声を聴かせ、本人の注意の向け方で神経活動パターンに違いがあるかどうか確かめるというものだ。

まず被験者にAさんとBさんの声をそれぞれ単独で聴かせ、側頭葉の神経活動パターンを画像として可視化する。次にAさんの声に注意した状態で、2人の声を合わせた音声を聴かせ、同じく可視化して見てみる。

結果、2人の声を単独で聴かせた場合は、聴覚野の活性化している部位や割合というものは双方で異なって表示された。また2人の声を合わせた

音声を聴かせた場合は、Aさん単独で聴かせた際の模様と似たような模様が表示された。

この実験から、聴覚野では注意の向け方に対応した神経活動、つまり聴覚情報の処理が起きていることがわかった。最近では、聴覚野の前段階である視床でも、注意の向け方に対応した神経活動があるのではないかと注目されている。注意の向け方が聴覚情報を処理する神経に与えている影響は非常に大きいと先生は話す。

これらの実験結果は、聞くという単純な行為が実は能動的な行為であるということ強く訴えている。聞き手の能動的な注意の向け方、つまりは聞き手の興味の方角によって、耳に入り認識される音は違ってくるのである。

さらにいえば聴覚とそれに連なる神経回路は、入ってきた音すべてを忠実に処理しているわけではない。その時々々の興味に応じて必要でなさそうなところは処理を飛ばし、必要そうなところは重点的に処理を行う。こういった処理のされ方の一つが、選択的聴取という例で浮き彫りになったと言えるのだ。

音声処理の問題としても、人間の選択的聴取の特性は注目されている。現代の音声処理技術は、話者となる音源の位置情報を頼りにすることで、音源相互の区別をしている。したがって、一つの場所から発せられた複数人の声について、話者の声を区別することは現代の技術では非常に難しい。一方で人間は話者の位置情報がなくとも、重複した音声から一人の声を拾い上げることができる。そこに既存の音声処理技術とは違った処理技術の原理の解明を期待する目が向けられている。

■ 音脈分凝

選択的聴取の例から明らかのように、一つのスピーカーから聞こえてくる複数人の声が重複した音声は、本来一つの流れでまとめて聞き取れるものではない。全員の人数はわからないにせよ、それぞれの話者の声を音の流れとして選択的に聴くことはできる。

声楽器に置き換わっても同じことが起こる。ロックサウンドであればギターとドラム、ベースなどの音は単一の音色さえわかっているならば、特定

の音色を聴き取ることが可能である。これはクラシックサウンドでも同様であり、音楽のジャンルを問わず同じことが言える。

一つの楽器の音色で構成された楽曲について、音の流れを区別することは可能だろうか。例えばオクターブを超える跳躍がある楽曲の旋律は、跳躍の直前と直後で別の流れとして解釈されるだろう。けれどもドとミの音など、音程が小さい場合は単なる音の上昇や下降の旋律として聞こえ、一つの流れとして解釈されると考えられる。

URL先のページにあるスライダーを動かすと繰り返し音の音が鳴り始める³⁾。より下へ動かしていくと、繰り返しの上下の音程がより大きくなる。

2の位置では、ソとラの2半音差の音が流れる。ソとラは一つの音の流れとして聞こえ、馬が駆けているような音に聞こえる。5の位置ではソとド#の6半音差の音が流れる。すると今度は流れが分かれたように聞こえ、下の音が細かいリズムで木魚を叩いているように聞こえる(図5)。この音を連続して聞いていると、途中で馬の駆ける音が

木魚を叩く音に変わったり、逆に木魚を叩く音が馬の駆ける音になったりする。この現象が、多義的知覚の一つの音脈分凝である。

この音の流れについての変化は、音に対する脳の解釈の変化を意味している。先生はこの例のような一定の音程の2つの音を規則的に鳴らしたときに、音の流れの解釈が変わるタイミングと、タイミングに前後して脳のどの部位が反応したかを調べる実験を行なった。ここから音脈分凝に関する脳のメカニズムを読み解くことができる。

先ほどの例の通り、一般に2半音差では1つの流れの方が聞きやすく、6半音差では2つの流れの方が聞きやすい。

つまり2半音差での音の流れが2つから1つになる変化および6半音差での音の流れが1つから2つになる変化は、聞きやすくなる向きの変化だと言えることができる。逆に、2半音差での音の流れが1つから2つになる変化および6半音差での音の流れが2つから1つになる変化は、聞きにくくなる向きの変化だと言えることができる。

実験した結果、聞きやすくなる向きの変化では、内側膝状体から聴覚野の順で反応が起こり、聞きにくくなる向きの変化では、聴覚野から内側膝状体の順で反応が起こった。つまり、聞きやすくなる向きの変化では、下位の内側膝状体が主導しており、聞きにくくなる向きの変化では、上位の聴覚野が主導していると言えることができる。

2000年代前半、音のまとまりの解釈は聴覚野およびそれ以降の大脳皮質が行なっているという考え方が、学会を始めとする専門家グループの主流の考え方であった。ところが、この柏野研究室の実験によって、内側膝状体が聴覚野と相互的に関わっていることが明らかになり、従来の考え方は大きく覆った。

2011年にも、より下位に位置する脳幹の下丘との関係を指摘する研究成果が外国の研究室によって報告されている。柏野研究室でも2013年、音のまとまりの知覚に対応した脳幹での神経活動を確認しており、聴覚野と脳のほかの部位とのさらなる関係性が指摘されている。

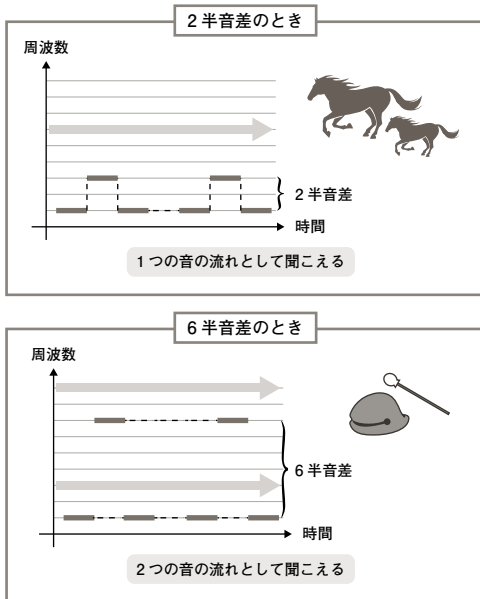


図5 音脈分凝の例

2半音差ではリズムカルに馬が駆けるような音、6半音差では規則的に打たれる木魚のような音が聞きやすい。

3) <http://www.brl.ntt.co.jp/illusionForum/a/frequencyDifference/ja/index.html> (※再生にはAdobe Flash Player環境を必要とする)

音楽の心地よさの研究

さて、脳が音楽の情報を無意識のうちに処理していると本稿の冒頭で述べたが、一体音楽はどのように情報として処理され、その結果として心地よさや魅力などを感じるのでしょうか。

音楽の心地よさといっても、個人や文化圏によって好むジャンルやテンポは異なる。また、楽器演奏などの音楽経験の有無や、その習熟度によっても同じ一つの楽曲に対する評価は変わってくる。私たちが日常的に耳にするような音楽に対して、いきなりその心地よさを判断するというのは困難である。

しかし、例えば子守唄と軍歌の曲調の違いは聴き手によらず歴然としているように、この二つを聞いた際の魅力や心地よさというものは大きく分けられる。このことから音楽の魅力や心地よさの中には、共通認識をもてるものがあると考えられることができる。

情動と不随意的な体の反応

心地よさといったものはまとめて、情動と呼ばれる。情動は、喜怒哀楽で喻えられる感情より基本的なものであり、快や不快、興奮や鎮静などに分類される。では、私たちが音楽を聴いているときの情動は、どのような経路を経ることで形成されるのだろうか。

音楽に関する聴覚情報は内耳を通り、脳幹や視床、聴覚野で処理される。処理された聴覚情報は、視床または聴覚野から扁桃体へと届けられる。扁桃体は情動に関する処理を行っており、聴覚情報は扁桃体を通して情動情報に変わる。

情動情報は自律神経の興奮やさまざまなホルモン分泌に繋がっている。例えば自律神経の一部をなす交感神経に伝わることで、発汗・動悸・震え・瞳孔の拡張などの体の反応が無意識的に起こる(図6)。体の反応自体が扁桃体に刺激を与えることもある。このように不随意的な体の反応というのは情動と密接に関わっている。ゆえに不随意的な体の反応を分析することで、気持ちの快や不快、興奮や鎮静といったものをある程度推定できると言える。

具体的な取り組み

先生は音楽の心地よさ、快や不快を評価するにあたってこの関係性を応用している。被験者に音楽を聴かせ、その人の瞳孔の大きさの変化をリアルタイムで計測することで、その人がどの程度快や不快を感じているのかを読み取ろうとする試みである。

しかし、瞳孔が拡張すれば快、収縮すれば不快という単純な関係にはならない。そこで先生はまず、被験者の主観に基づく快や不快についてのリアルタイムでの評価と、瞳孔の大きさの変化を比較する実験を行なった。電子機器のダイヤルを被験者に触れさせ、快だと感じたら左に、不快だと感じたら右に回してもらい、これを主観に基づく評価とした。ここから瞳孔の大きさの変化に対応する被験者の快や不快を推定することができた。

音楽を聴くことによって生じる快や不快を評価するためには、音楽に対する快や不快についての客観的な指標が必要になる。この指標を決めるにあたって、先生は聴いている人が予想する曲の展開と実際の曲の展開とのギャップに着目した。旋律、曲調、テンポなどが新たな展開に入ったとき、人はその前後のギャップに快感を覚えたり、逆に不快に感じたりする。先生はこのギャップの大きさをサプライズと呼び、楽曲のリアルタイムでの展開の度合いを指標とした。

何に基づいて計算するかによっていろいろなサプライズを決めることができる。音程の時々刻々の変化から、各楽章の曲色の変化に至るまで、多

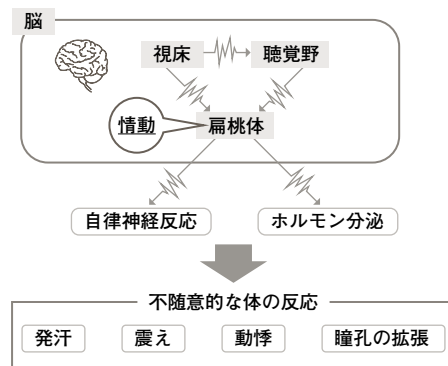


図6 扁桃体と情動情報の流れ

様な規模でサプライズを定義することができる。

具体的に先生は、ベイズ推定と呼ばれる統計学の技法を用いた(図7)。楽曲の展開における t 秒後と $t + \Delta t$ 秒後のギャップを数学的に評価した。音楽が始まって0秒後から t 秒後までに登場する音の周波数および頻度を確率分布の曲線で表し、それを $t + \Delta t$ 秒後における曲線と比較する。この差が大きいものをサプライズが大きい、差が小さいものをサプライズが小さいと評価した。

実際に先生は、ベイズ推定で定めた音楽側の指標と、情動と結びついている瞳孔の大きさの変化との関係性を探っている。先生はさらに良い対応を得るべく別のサプライズの定義を考案し、検討を進めている。

研究の近況と展望

近年、柏野先生は自閉症スペクトラム障害の患者に関心をもち、被験者として研究の対象にしている。自閉症スペクトラム障害とは、発達障害の一種で、ここ数年、脳や遺伝子などの観点から急速に研究が進んでいる。知能の良し悪しには関係なく、一般に社会性やコミュニケーション能力の獲得などの発達の遅れを特徴とする障害である。

数年前、聴覚の異常で先生のもとを訪ねてきた自閉症スペクトラム障害の女性がいた。彼女は、さほど混雑していない喫茶店程度の騒がしさの中でさえ、会話することができないことを訴えていた。騒音が無い二人だけの環境で会話することはできるが、そこに他人の声や雑音が入っただけで、話している相手の声がかき消されてしまう、というのだ。彼女は、勤め先でもこの問題に直面しており、仕事内容について上司の指示が聞こえないということで上司から注意を受けていた。何回も耳鼻科へ掛かりにいったが、聴覚能力には異常がないと診断された。

ここでの聴覚能力に異常がないとは、高い音・低い音が聞き取れるかどうかという純音測定の結果であり、耳鼻科で一般的に行われるテストで確認されるものだ。彼女の場合、単独の音が聞こえるかどうかという点では問題は見られなかった。

では、彼女の聴覚に関してどこに具体的な異常

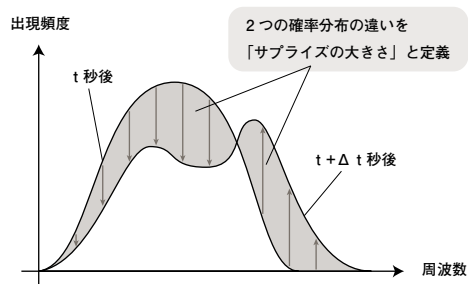


図7 ベイズ推定

があったのだろうか。先生は、自閉症スペクトラム障害では、選択的聴取に関わる神経機構が先天的に欠損している場合があることを指摘する。選択的聴取の能力をもっていないがために、会話する相手の声を耳に入ってくる音声から取り出すことができないのではないかと、ということだ。

自閉症スペクトラム障害の人の聴覚情報の処理に関わる脳幹の細胞が異常な配列をしていたという研究成果も出てきている。これは例の女性の聴覚特性に見受けられたような、自閉症スペクトラム障害の人がもつ独特な感覚特性が、情報処理回路自体の異常によって説明される根拠となる。さらに自閉症スペクトラム障害自体も、こういった情報処理回路の異常によって説明される可能性がある」と先生は話す。

人間の脳は物事すべてに対して意識的に判断できているわけではない。加えて意識的な判断は、無意識下のさまざまな固定概念に縛られている。錯覚の例に見たように無意識による制約は人間の不正確さとして映るが、その無意識的な脳のはたらきがあってこそ支障なく日常生活を送っていくことができるのである。

先生は脳の無意識的なはたらきに着目し、認知科学、脳科学などの多様な学問領域から、これからも人間の感覚特性を明らかにしていく。

執筆者より

先生が話してくださった内容はどれも私自身の、人間の感覚や認識にまつわる根本的な興味に訴えるものがありました。ありがとうございました。

(小沢陽)