

論文 / 著書情報  
Article / Book Information

|                   |                                                                                                                                                                                       |
|-------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 題目(和文)            | 限られた入力からの学習を通じた聴覚機能の獲得: 人工ニューラルネットワークを用いた計算モデリングの研究                                                                                                                                   |
| Title(English)    | Acquisition of Auditory Functions through Learning from Limited Input: A Computational Modeling Study Using Artificial Neural Networks                                                |
| 著者(和文)            | 芦原孝典                                                                                                                                                                                  |
| Author(English)   | Takanori Ashihara                                                                                                                                                                     |
| 出典(和文)            | 学位:博士(理学),<br>学位授与機関:東京科学大学,<br>報告番号:甲第3号,<br>授与年月日:2024年12月31日,<br>学位の種別:課程博士,<br>審査員:金子 寛彦,中村 健太郎,小池 康晴,吉村 奈津江,小尾 高史,柏野 牧夫                                                          |
| Citation(English) | Degree:Doctor (Science),<br>Conferring organization: Institute of Science Tokyo,<br>Report number:甲第3号,<br>Conferred date:2024/12/31,<br>Degree Type:Course doctor,<br>Examiner:,,,,, |
| 学位種別(和文)          | 博士論文                                                                                                                                                                                  |
| Category(English) | Doctoral Thesis                                                                                                                                                                       |
| 種別(和文)            | 論文要旨                                                                                                                                                                                  |
| Type(English)     | Summary                                                                                                                                                                               |

# 論文要旨

## THESIS SUMMARY

|                                  |             |     |                           |           |        |
|----------------------------------|-------------|-----|---------------------------|-----------|--------|
| 系・コース：                           | 情報通信        | 系   | 申請学位 (専攻分野)：              | 博士        | ( 理学 ) |
| Department of, Graduate major in | ライフエンジニアリング | コース | Academic Degree Requested | Doctor of |        |
| 学生氏名：                            | 芦原 孝典       |     | 審査員主査：                    | 金子 寛彦     |        |
| Student's Name                   |             |     | Chief Examiner            |           |        |

### 要旨 (和文 2000 字程度)

Thesis Summary (approx.2000 Japanese Characters )

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>本論文は「Acquisition of Auditory Functions through Learning from Limited Input: A Computational Modeling Study Using Artificial Neural Networks (限られた入力からの学習を通じた聴覚機能の獲得: 人工ニューラルネットワークを用いた計算モデリングの研究)」と題し、全4章で構成される。</p> <p>第1章「Introduction (序論)」では、本研究の背景について述べる。聴覚において興味深い点の1つは、連続聴効果や音脈分凝といった錯聴に代表されるように、刺激の物理的特性と知覚内容が単純には対応しない点である。この現象が生得的なものか、学習によるものかについてはまだ不明な部分も多いが、神経可塑性による処理の変容や知覚学習の効果を考慮すると、学習が知覚形成に関与している可能性は極めて高い。そこで本論文ではこの学習の効果や限界を明らかにすることを目的としている。そして、2つのテーマを通じて、どの程度の知覚・認知機能が学習によって獲得可能かを検討した。第一のテーマ(第2章)では、聴覚刺激は不完全だが対応する教師が明確に存在する場合として、時間・スペクトル情報が制限された人工内耳信号から基本周波数(F0)が推定可能かを検討した。第二のテーマ(第3章)では、聴覚刺激に顕著な劣化がないが、対応する教師が明確に存在しない場合として、音声刺激だけで言語知識が獲得可能かを検討した。</p> <p>第2章「Estimating Pitch Information from Simulated Cochlear Implant Signals with Deep Neural Networks (ディープニューラルネットワークによる人工内耳模擬信号を用いたピッチ推定)」では、比較的低次の処理としてピッチ知覚に着目する。人工内耳とは、重度の難聴を回復させる人工臓器である。一般的に人工内耳を装着することで音声知覚は大きく回復するが、ピッチ知覚が困難であると報告されている。しかし、その原因が人工内耳の信号によるものか、それとも他の要因(装着時期・期間や聴神経生存率など)によるものかは明らかではない。その点を明らかにするため、本章では音波形において高精度にF0推定が可能な人工ニューラルネットワーク(ANN)を導入し、その聴覚刺激を人工内耳信号に置き換えた場合のF0推定精度を計測した。特に人工内耳信号は、パルスレート(時間解像度)とチャンネル数(スペクトル解像度)を系統的に調整出来るため、学習による効果・限界を探索するのに適している。実験結果から、静音環境下でかつ一定のチャンネル数・パルスレートが確保されていれば、音波入力と遜色ない精度でF0推定が可能であることが明らかとなり、それは人工内耳信号には、F0推定に必要な情報が一定量含まれていることを示している。この結果は、大きく劣化した聴覚刺激でも、学習によってピッチ推定がある程度可能であることを示している。</p> <p>第3章「Unveiling the Linguistic Capabilities of a Self-Supervised Speech Model through Cross-Lingual Benchmark and Layer-Wise Similarity Analysis (クロスリンガルベンチマークと類似度分析による自己教師あり学習モデルの言語知識分析)」では、比較的高次の処理として音声言語獲得に着目する。音声言語知識の獲得には、一般的にジェスチャーやテキストなどの情報が不可欠と考えられる一方、胎児や乳幼児の段階で音声だけから様々な情報処理基盤が形成される可能性も示唆されている。そこで本章では、教師が明確に存在しない場合として、音声刺激だけから学習を通じてどの程度の言語知識が獲得可能かを明らかにすることを目的としている。具体的には、自己教師あり学習(SSL)と呼ばれる教師なし学習法の一つを用いて構築されたモデル(音声SSLモデル)を用いて検討を行なった。このモデルに対し、文の意味や構文知識を評価項目とするベンチマークを用いて、基本的な音声信号処理に基づくベースラインおよび、テキストだけで学習されたSSLモデルと精度を比較した。実験結果から、音声SSLモデルはベースラインよりも高い精度を示したが、テキストSSLモデルよりも低い精度にとどまった。このことから、テキスト情報は言語知識の獲得において重要であることを示すとともに、聴覚刺激だけでも学習を通じて一定の言語知識を獲得可能であることを明らかにした。</p> <p>第4章「Conclusions and Future Work (結論と今後の課題)」では、本論文の結論と今後の課題について述べる。以上を要するに、本論文では、限られた入力であっても、聴覚刺激さえ与えられれば、学習を通じて、様々な知覚・認知機能を獲得できる可能性をモデルによる検討より明確に示した。具体的には、例えば先天性難聴であっても、補聴器や人工内耳を通じて適切な聴覚刺激を提供し、健常者と同等の教育・学習過程を経験させることで、健常者と同等の聴覚・認知機能を獲得出来る可能性がある。今後、これらの人工ニューラルネットワークによる計算モデリングによる知見が、人間の聴覚情報処理機構を解明する上での一つの基本的指標となることを期待する。</p> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東京科学大学リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Science Tokyo Research Repository Website (T2R2).

## 論文要旨

### THESIS SUMMARY

|                                  |             |     |                           |           |      |
|----------------------------------|-------------|-----|---------------------------|-----------|------|
| 系・コース：                           | 情報通信        | 系   | 申請学位(専攻分野)：               | 博士        | (理学) |
| Department of, Graduate major in | ライフエンジニアリング | コース | Academic Degree Requested | Doctor of |      |
| 学生氏名：                            | 芦原 孝典       |     | 審査員主査：                    | 金子 寛彦     |      |
| Student's Name                   |             |     | Chief Examiner            |           |      |

### 要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words)

This thesis explores whether auditory functions can be acquired through learning from limited inputs using a computational modeling approach.

Chapter 1 provides the general background for this thesis. One intriguing aspect of auditory perception is that the physical properties of stimuli do not straightforwardly correspond to perceptual content, suggesting that learning from auditory stimuli would play a significant role. This thesis focuses on the effects and limitations of learning, which typically requires auditory stimuli and corresponding teachers. The first theme (Chapter 2) investigates a scenario where auditory stimuli are incomplete, but a corresponding teacher is clearly present. The second theme (Chapter 3) explores a situation where the auditory stimuli are not significantly degraded, but there is no clear corresponding teacher.

Chapter 2 investigates the feasibility of F0 estimation from cochlear implant (CI) signals, which convey limited information due to device restrictions. CI users, even those with substantial speech comprehension, typically have poor sensitivity to pitch information (or fundamental frequency, F0). To explore whether this difficulty can be attributed to CI signals, we estimate the F0 information in CI signals using an artificial neural network (ANN). The results indicate that, with sufficient channels and pulse rates in quiet environments, the performance is comparable to that of waveform input, suggesting that CI signals contain a certain level of F0 information.

Chapter 3 explores the feasibility of spoken language acquisition from speech stimuli alone. While multimodal information (e.g., gestures and text) is typically considered necessary for speech-language development, several studies have suggested that the information-processing capabilities of a fetus or infant may be formed solely from auditory input. This chapter examines the language acquisition capabilities of an ANN model trained on speech inputs alone by performing various spoken language understanding tasks that require semantic and syntactic knowledge. The results show that the model's performance surpasses that of a baseline based on basic speech signal processing, indicating that language knowledge can be acquired from speech input alone.

Chapter 4 presents the conclusions and future directions of this thesis.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東京科学大学リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Science Tokyo Research Repository Website (T2R2).