

聴覚刺激列の統計学習の神経生理学的研究

湯本 真人¹⁾、大黒 達也²⁾

Key words: 統計学習、単語分節化、統語、遷移確率、マルコフ連鎖、脳磁図

【要旨】聴覚領域における統計学習は、音声言語の単語分節化として行動指標に基づいて心理学的に研究されることが多く、神経生理学的研究は数少なかった。単語分節化の研究で広く用いられる3音節疑似単語のランダム連結は、単純マルコフ連鎖と見做すことができる。より一般化された高次の確率過程で規定される刺激系列を用いて、統計学習を神経生理学的指標で評価する手法により、これまで余り光が当てられていなかったヒト脳の学習機構の方略に迫れる可能性がある。本稿では、第22回認知神経科学会教育講演Ⅱにおける講演の内容から、既に学術論文にて発表済みのものを概説した。

はじめに

一連の新奇な刺激に出会ったとき、人は刺激の順列パターンや分布を、その中に含まれる統計的な構造を抽出しながら学習する。この機構は統計学習とよばれ、人が生来的に持っている基本的で領域普遍的な学習プロセスと考えられている。聴覚領域における統計学習の研究が脚光を浴びた研究として、Saffran JRらの研究が挙げられる¹⁾。無意味単語“bidaku”, “padoti”, “golabu”などを切れ目なく連続させた合成音声で短時間聞かせただけで、8か月の乳児でも単語の切り出し (word segmentation; 単語分節化) が可能であることを示した。単語内の音節の遷移 (bi → da や do → ti など) は100%の確率で起きるのに対し、単語間の音節の遷移 (ku → pa や bu → bi など) は、語順に自由度がある分、生起確率としては低くなる。この単語内・単語間の音節遷移確率の違いを学習することで単語分節化が行われるものと考えられ、音声言語の獲得過程における統計学習の例として注目を集めた。聴覚領域における統計学習の研究は、乳児の単語分節化課題を中心として、行動指標による心理実験が主流であった。

聴覚統計学習を神経生理学的指標で捕えようとする研究も、単語分節化課題を用いて行われている^{2,3)}。三音節からなる無意味単語を、同じ単語が連続しないよう統制をかけた疑似ランダムな順列で結合した連続音を提示し、単語内・単語間の音節遷移確率の不均等分布を学習することにより単語分節化が達成された結果、単語の冒頭の音節に対する事象関連電位がその他の音節に対する電位に比べて増高する。また、非言語音の三音の順列を単語 (tone word) と見做して、単語分節化課題同様に複数の tone word を切れ目なく繋げた連続音として提示しても、分節化による事象関連電位上の変化が捕えられ⁴⁾、脳磁場応答の変化としても捕えられている⁵⁾。

このように、多くの聴覚統計学習の神経生理学的研究は、tripletを単語とした単語分節化課題を用いて行われてきているが、単語内の音の遷移確率は100%で確定し、単語間はランダムに配列されている構造上、事象関連電位に内包される情報は単語分節化の程度に限定される。また、単語がtripletで構成されているため、一旦tripletが単位であることが学習されると、その構造に対する知識がそれ以降の学習においてcueとなり得るなど制約もある。単語分節化課題は統計学習を研究す

¹⁾ 東京大学大学院医学系研究科病態診断医学講座

²⁾ マックス・プランク認知脳科学研究所

る上で一つの特例形に過ぎず、より一般化された課題を用いることにより、これまでの統計学習の研究では光が当てられることの無かった側面を明らかにできるのではないかと考えられる。

I. マルコフ連鎖

事象関連応答の MMN (mismatch negativity) や P300 を記録する際の刺激提示法として広く用いられているオドボール系列の最も単純な様式は、高頻度に出現する標準刺激 s に混じって標的刺激 d が低頻度で出現する。低頻度の標的刺激が 2 回連続して出現しないように統制をかけた疑似ランダム系列の状態遷移図を図 1a に示す。直前の刺激によって規定される遷移確率で次の刺激が出現することを示しており、疑似ランダム単純オドボール系列は単純マルコフ連鎖⁶⁾ と見做すことができる。

単語分節化課題の状態遷移図を図 1b に示す。こちら、直前に提示した音節によって確率 1 または $1/(n-1)$ で次の音節が決まるため、単純マルコフ連鎖の特例形である。単語や文節などの文法構造を模したいわゆる人工文法系列も、不均等な遷移確率分布で同様に表される単純マルコフ連鎖を用いたものが多い。

このように、確率的な状態遷移として刺激系列を捉えると、これまでの研究で用いられている確率過程はやや特殊例であることから、より一般形に近い確率過程を刺激系列に用いて、一から、ヒトの統計学習を行動指標と神経生理学的指標で押さえて行きたい。一例として、直近の 2 つの刺激の順列で次に遷移する刺激が一様な確率分布で規定される二重マルコフ連鎖の例を図 1c に示す。

このような連鎖を、例えば高頻度遷移確率 0.8、低頻度遷移確率 0.05 (低頻度遷移刺激 4 種で $0.05 \times 4 = 0.2$) で順列提示した場合、高頻度遷移が 3 回連続する確率は $0.8^3 = 0.512$ でほぼ 50%、高頻度遷移が連続する回数の期待値は、級数 $\sum_{n=1}^{\infty} (n \times 0.8^n \times 0.2 = 4)$ となる (一般に高頻度遷移確率が p

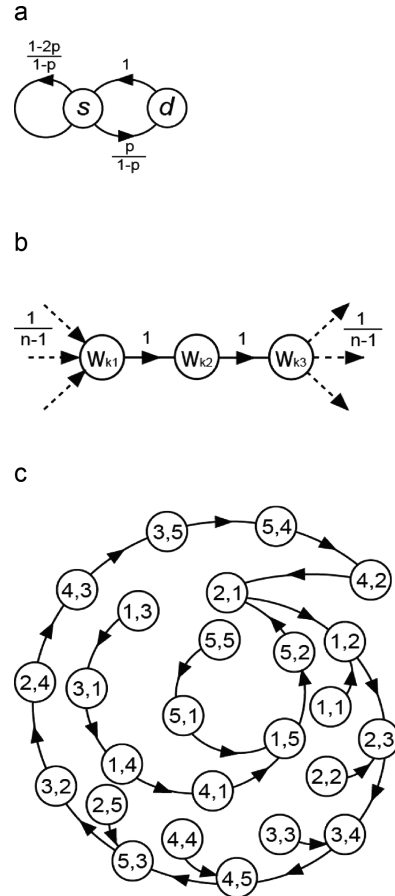


図 1 各種刺激系列の状態遷移図 (文献⁷⁾ より改変)

a. 疑似ランダム単純オドボール系列
標的刺激 d の出現確率を $p (< 0.5)$ とすると $P(s) + P(d) = 1$. 疑似ランダム ($P(d|d) = 0, P(s|d) = 1$) より $P(d) = P(d|s)P(s) + P(d|d)P(d) = P(d|s)P(s) = p, P(s) = P(s|s)P(s) + P(s|d)P(d) = P(s|s)P(s) + P(d)$. よって、 $P(d|s) = p/(1-p), P(s|s) = (1-2p)/(1-p)$.

b. 単語分節化系列

n : 単語数、 W_{k1} : 単語 k の第 1 音節、 W_{k2} : 単語 k の第 2 音節、 W_{k3} : 単語 k の第 3 音節 ($1 \leq k \leq n$). 単語 k 以外の $n-1$ の単語の第 3 音節から確率 $1/(n-1)$ で W_{k1} に遷移。以後 $W_{k2} \rightarrow W_{k3}$ と確率 100% で遷移し、単語 k 以外の $n-1$ の単語の第 1 音節に確率 $1/(n-1)$ で遷移する。

c. 二重マルコフ連鎖に基づく刺激系列

○内は 5 種類の刺激 1~5 の最近の順列を示し、矢印は高頻度で遷移する次の状態を示す。例えば、(5,1) は刺激 5 → 刺激 1 の順に刺激提示された状態を示し、矢印の指す (1,5) が高頻度で遷移する次の状態 (即ち次の刺激は 5) を示す。5 以外の刺激 1, 2, 3, 4 は低頻度で提示されるが、低頻度遷移を示す矢印は省略されている。一般に、最近の n 個の順列事象によって次の事象が確率的に統制される連鎖を、 n 重マルコフ連鎖と呼ぶ ($n > 1$; $n = 1$ の場合は単純マルコフ連鎖)。

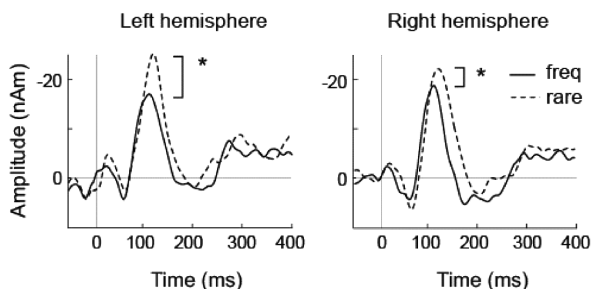


図2 学習による聴覚誘発脳磁場応答の変化 (文献 (Daikoku et al. 2015) Fig. 6h より)

なら同期期待値は $p/(1-p)$ 。平均4回の高確率遷移連鎖を万遍なく断片的に提示される間に、被験者は頻繁に出現する刺激の連鎖を学習し、低頻度遷移で出現する刺激に対する応答に比べ、高頻度遷移で出現する刺激に対する応答が減弱していくと予想される。

実際に脳磁場応答でこのことを確認した結果を図2に示す。高頻度遷移音に対する応答 (freq) と低頻度遷移音に対する応答 (rare) は、同一の刺激アンサンブルに対する応答の加算平均となるよう統制をかけているため、聴取開始時には原理的に差異はないはずだが、一定時間経過後、freqはrareに比べて有意に減衰する。この差は、音列の遷移確率の差によるものと考えられ、聴取の間に遷移確率の統計学習が行われた結果生じたものと言える。

統計学習を神経生理学的指標で捕えた研究は、行動指標に基づく研究に比べ数少ない⁸⁾。統計学習の前後で生じる差波形はMMNに類似しており、Statistical MMNと呼ぶ研究者もいる⁹⁾。マルコフ連鎖に埋め込まれた確率遷移のオドボールは、一定の学習を達成した後でオドボールとして機能する点で、先に示した疑似ランダム単純オドボール系列の場合とは若干背景が異なる。MMNは古典的には、直前の刺激の感覚記憶痕跡との照合過程を反映していると考えられているが、二重マルコフ連鎖の例では、より時間的に離れた事象の記憶との照合過程を反映していると考えられ、

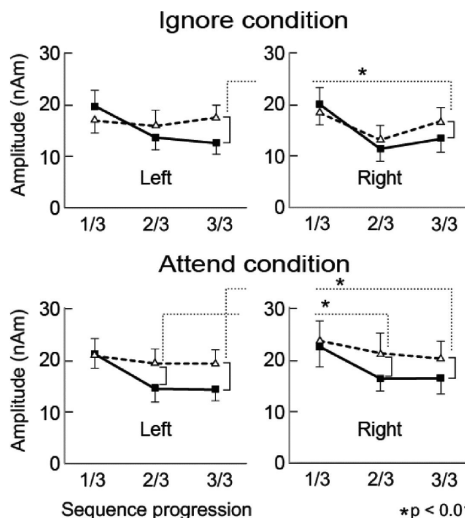


図3 無視条件 (上段)、注意条件 (下段) における聴覚誘発脳磁場応答 N1m の振幅の推移 (聴取シーケンスの進行 (1/3 → 2/3 → 3/3) に伴い、注意条件では2/3の段階で高頻度遷移音に対する応答 (実線) は低頻度遷移音に対する応答 (破線) よりも有意に低下しているのに対し、無視条件では3/3の段階で有意差が検出された。 * $p < 0.01$)

MMN生成機構よりも長い時間保持される記憶の関与が想定される。加えて、刺激から潜時100ms以降の成分のみでなく、潜時50ms付近の成分 (P1やP1m)にも学習効果による高確率遷移に対する応答の減弱が報告されている^{5,10)}点も、古典的MMNとは様相を異にしている。

II. 注意の影響

統計学習は一般に、学習対象に対する知識を持たず、また学習していることを意識しない形で進む潜在学習である。複合音による二重マルコフ連鎖を刺激系列として、刺激に対する注意が統計学習に及ぼす影響を調べた結果、無視条件よりも注意条件の方が、学習効果 (高頻度遷移音に対する応答の相対的低下) が早期に表れることが示された (図3)¹¹⁾。

また、それぞれマルコフ連鎖で構成される高・低二つの音列を同時に聴取させる課題において、一方に注意を向けた場合と両者に注意を向けた場合とで、それぞれの音列の統計学習の成績を比較

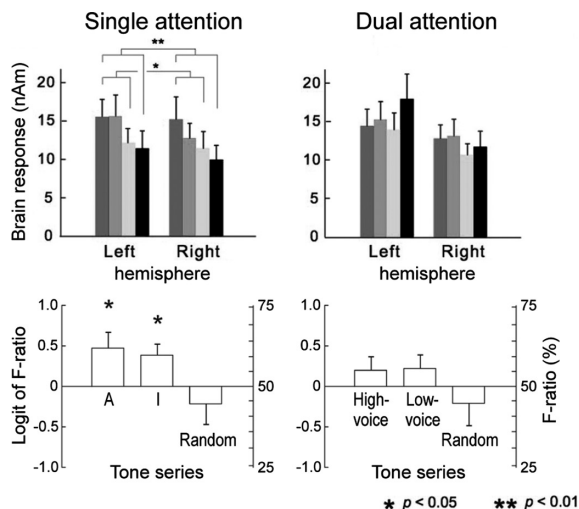


図4 一方の音列にのみ注意を向けた場合（左）、両方に注意を向けた場合（右）の聴覚誘発脳磁場応答 P1m の振幅（上段）と聞き覚えテストの結果（下段）

一方の音列のみに注意を向けた場合、双方の音列とも低頻度遷移音だった場合の脳活動（赤）、注意を向けた音列が低頻度遷移音で無視した音列が高頻度遷移音だった場合の脳活動（緑）、注意を向けた音列が高頻度遷移音で無視した音列が低頻度遷移音だった場合の脳活動（黄）、双方の音列とも高頻度遷移音だった場合の脳活動（青）、の順に強い脳活動を認め（左上）、脳活動の記録後に行った聞き覚えテストでも、注意を向けた音列（A）だけでなく無視した音列（I）の断片に対しても有意に「聞き覚えがある」との回答が得られた（左下）。一方、両方の音列に注意を向けた場合は、脳活動に有意差を認めず（右上）、音列断片に対して有意に「聞き覚えがある」との回答は得られなかった（右下）。

したところ、一方に注意を向けた場合に、無視した音列の学習効果も含めて学習効果が上がることが示された（図4）¹²⁾。

III. 言語と音楽

マルコフ連鎖刺激系列において使用する刺激音の音響特性を変化させることで、言語的・音楽的な音列の統計学習の違いを調べることができる¹³⁾。無意味単語の順列を二重マルコフ連鎖で確率的に統制することにより、語順の学習が先に検出され、その後に単語分節化が検出されることが示された¹⁴⁾。聴覚言語の統計学習が、統語から単語へと進む可能性を示唆している。単語分節化は単語の順列に多様性が無ければ起き得ないが、単語を疑似ランダムに並べた単語分節化系列による行動指標に基づくこれまでの研究手法では、単語分節化と語順のような異なるレベルの学習の時間関係を調べることは困難であった。刺激系列に確率的構造を持ち込み、行動指標だけでなく神経生理学的指標に基づいて統計学習を評価する手法の利点と言える。また、途中で語順の規則を変更することにより、獲得した統語レベルの知識を修正するには、新規の学習よりも時間がかかることが示唆された。

音楽的な音列として、和音進行の研究がこれまでも行われているが、我々の生育環境には西洋音楽が溢れているため、結果の解釈には、母体内に始まる潜在学習により獲得された音楽的知識の影響を排除できない。そこで、12平均律や機能と和声から乖離した極力聞き覚えのない和音列を用いて和音進行の統計学習を調べたところ、無調和音列の統計学習にはピッチクラスの影響が関与することが示され、和音列の統計学習はP1振幅に反映することが示唆された¹⁰⁾。

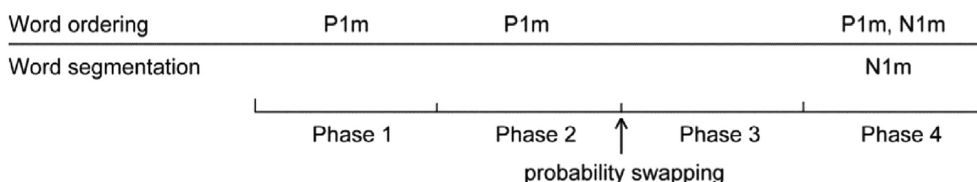


図5 語順・単語分節化の統計学習の誘発磁場成分への反映の時間経過
語順の学習は早期（Phase 1 & 2）から P1m の振幅差として検出され、語順規則の遷移確率の反転後、Phase 4 で再び語順学習が P1m, N1m 振幅差として検出され、同時に初めて単語分節化の学習も N1m 振幅差として検出された。

おわりに

これまで統計学習は行動指標に基づく心理実験により研究されることが圧倒的に多く、神経生理学的研究は数少なかった。刺激系列に注目すると、疑似ランダム単純オドボール系列や単語分節化系列は単純マルコフ連鎖と見做すことができる。より一般化された確率過程で規定される刺激系列を用いて確率統計構造を操作し、神経生理学的指標により統計学習効果を検出する手法により、これまで余り光が当てられていなかったヒト脳の学習機構の方略に迫れる可能性がある。新たな興味深い研究成果が公表されつつある中、本稿では、第22回認知神経科学会教育講演IIにおける講演の内容から、既に学術論文にて発表済みのものを概説した。

文献

- 1) Saffran JR, Aslin RN, Newport EL. (1996) Statistical learning by 8-month-old infants. *Science* 274(5294), 1926-1928.
- 2) Sanders LD, Newport EL, Neville HJ. (2002) Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat Neurosci* 5(7), 700-703. doi: 10.1038/nn873
- 3) Teinonen T, Fellman V, Naatanen R, Alku P, Huotilainen M. (2009) Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci* 10, 21. doi: 10.1186/1471-2202-10-21
- 4) Abia D, Katahira K, Okanoya K. (2008) On-line Assessment of Statistical Learning by Event-related Potentials. *J Cogn Neurosci* 20(6), 952-964. doi: 10.1162/jocn.2008.20058
- 5) Paraskevopoulos E, Kuchenbuch A, Herholz SC, Pantev C. (2012) Statistical learning effects in musicians and non-musicians: an MEG study. *Neuropsychologia* 50(2), 341-349. doi: 10.1016/j.neuropsychologia.2011.12.007
- 6) Markov AA. (1971) Extension of the limit theorems of probability theory to a sum of variables connected in a chain. Reprinted in Appendix B of: Howard, R. *Dynamic Probabilistic Systems, Volume 1: Markov Chains*. John Wiley and Sons.
- 7) Yumoto M, Daikoku T. (2016) Part IV Auditory System: Basic Function. In: *Clinical Applications of Magneto encephalography* (Tobimatsu S, Kakigi R, editors). Springer Japan; 2016. pp. 97-112.
- 8) Furl N, Kumar S, Alter K, Durrant S, Shawe-Taylor J, Griffiths TD. (2011) Neural prediction of higher-order auditory sequence statistics. *Neuroimage* 54(3), 2267-2277. doi: 10.1016/j.neuroimage.2010.10.038
- 9) Koelsch S, Busch T, Jentschke S, Rohrmeier M. (2016) Under the hood of statistical learning: A statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. *Sci Rep* 6: 19741. doi: 10.1038/srep19741
- 10) Daikoku T, Yatomi Y, Yumoto M. (2016) Pitch-class distribution modulates the statistical learning of atonal chord sequences. *Brain Cogn* 108: 1-10. doi: 10.1016/j.bandc.2016.06.008
- 11) Daikoku T, Yatomi Y, Yumoto M. (2014) Implicit and explicit statistical learning of tone sequences across spectral shifts. *Neuropsychologia* 63, 194-204. doi: 10.1016/j.neuropsychologia.2014.08.028
- 12) Daikoku T, Yumoto M. (2017) Single, but not dual, attention facilitates statistical learning of two concurrent auditory sequences. *Sci Rep* 7(1), 10108. doi: 10.1038/s41598-017-10476-x
- 13) Daikoku T, Yatomi Y, Yumoto M. (2015) Statistical learning of music- and language-like sequences and tolerance for spectral shifts. *Neurobiol Learn Mem* 118, 8-19. doi: 10.1016/j.nlm.2014.11.001
- 14) Daikoku T, Yatomi Y, Yumoto M. (2017) Statistical learning of an auditory sequence and reorganization of acquired knowledge: A time course of word segmentation and ordering. *Neuropsychologia* 95, 1-10. doi: 10.1016/j.neuropsychologia.2016.12.006

Neurophysiological Studies on Auditory Statistical Learning

Masato Yumoto¹⁾, Tatsuya Daikoku²⁾

¹⁾Department of Clinical Laboratory, Graduate School of Medicine, The University of Tokyo

²⁾Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences

Key words : statistical learning, word segmentation, syntax, transition probability, Markov chain, magnetoencephalography

Statistical learning in auditory domain has been predominantly investigated as an effect of word segmentation in speech sequences by psychological studies based on behavioral measures, and few neurophysiological studies have been conducted. Random concatenation of pseudo-words that consist of three syllables, which has been widely used to study word segmentation, can be regarded as a simple Markov chain. By introducing more generalized and higher-order stochastic processes into stimulus sequence, neurophysiological measure of statistical learning may potentially shed light on the new aspect of human learning strategy. In this paper, recent findings that have been published are briefly reviewed.