計算論的神経心理学

浅川伸一

ここで取り上げるモデルは入出力モダリティによって以下のように分類できる:

1. **読字モデル** 三角89 (Seidenberg & McClelland, 1989), 三角96 (Plaut, McClelland, Seidenberg, & Patterson, 1996)，DRC (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001)，CDP++ (Perry, Ziegler, & Zorzi, 2010; Zorzi, Houghton, & Butterworth, 1998), 交差点モデル (Kello, 2006)，系列符号化モデル (Sibley, Kello, Plaut, & Elman, 2008)
2. **命名モデル** SP モデル (Foygel & Dell, 2000; Hickok, 2012), SLAM (Walker & Hickok, 2016), MPT (Walker, Hickok, & Fridriksson, 2018)
3. **聴入力モデル** Lichtheim2 (Kello, 2006), WEAVER (Levelt, Roelofs, & Meyer, 1999; Roelofs, 1997, 2014)
4. **音韻＝発話モデル** DIVA (Tourville & Guenther, 2011), GODIVA (Guenther, 2016), C/D (Fujimura, 1992, 2000, 2003)

読字モデルは Norris (2013) の総説論文によれば，初期のモデルは，オモチャモデルと呼ばれる制限されたモデルであっpた。現実的な語彙数に向けてモデルが発展してきている。 (Norris, 2013, Tab. 1) によれば，読字過程のモデルとして，CDP++ (Perry et al., 2010), DRC (Coltheart et al., 2001), 三角モデル (Plaut et al., 1996; Seidenberg & McClelland, 1989), 系列符号化モデル (Sibley et al., 2008), 交差点モデル(junction) (Kello, 2006) が挙げられている。

読字モデルの入力を詳細に検討すれば相互活性化モデル interactive activation models McClelland & Rumelhart (1981); Rumelhart & McClelland (1982) に端を発する。相互活性化モデルをやや詳しく解説する。

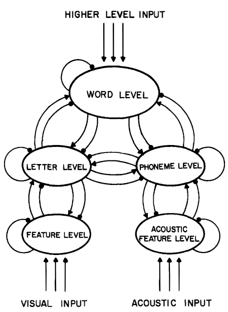


Figure 1. A sketch of some of the processing levels in-volved in visual and auditory word perception, with in-terconnections.

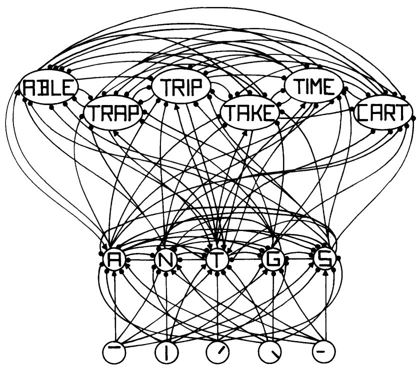


Figure 3. A few of the neighbors of the node for the letter T in the first position in a word, and their interconnections.

図1　McClelland & Rumelhart (1981) より

静止膜電位 r_i は先見バイアスに相当する (Broadbent, 1967) 長期間の活性状態を反映し，高頻度語では静止膜電位レベルは高く，低頻度語では低い。現代的な言葉では，まさにバイアスである。

時刻 t におけるノード i の活性値を $a_i(t)$ と表記する。そのノードの静止膜電位レベルを r_i とする。静止膜電位への回帰減衰率を Θ_i とする。ノード i への全入力を $n_i(t)$ と表記すれば，次式で表すものとする:

$$n_i(t) = \sum_j \alpha_{ij} e_j(t) - \sum_k \gamma_{ik} i(t), \tag{1}$$

ここで $e_j(t)$ は正の入力であり，$i_k(t)$ は負の入力を表すものとする。 α_{ij} および γ_{ik} は結合係数である。現代的に表記すれば $w_{ij} \in \mathbb{R}$ とすれば，$n_i(t) = \sum_j^J w_{ij} x_j$ のような表記となる。

任意のノードの入力の影響は，当該ノードの状態と最大活性値との差分に応じて増幅される:

$$\epsilon_i(t) = n_i(t)\left(M - a_i(t)\right), \tag{2}$$

ここで M そのノードの最大活性値である。式(2) を目標関数にして学習させるとどうなるのか

Δt 時刻後の活性値 $a_i(t+\Delta t)$ は減衰関数 Θ と近傍ノードの影響 ϵ_i とによって次式で表される:

$$a_i(t+\Delta t) = a_i(t) - \Theta_i\left(a_i(t) - r_i\right) + \epsilon_i(t). \tag{3}$$

一方で，

$$e_i(t) = a_i(t)\left(a_i(t) - m\right), \tag{4}$$

ここで m は最小活性値である。反応について:

$$\bar{a}_i(t) = \int_{-\infty}^{t} a_i(x)\, e^{-(t-x)r} dx. \tag{5}$$

(5) RHS において x は時刻ではり $[-\infty, t]$ の範囲をとるものとする。指数関数はそのノードの活性値の影響を表すものとする。パラメータ r は新旧刺激の貢献度合いを定めるパラメータである。

反応強度 s_i は Luce のモデルに従って次式で与えられる:

$$s_i(t) = e^{\mu \bar{a}_i(t)}, \tag{6}$$

ここでパラメータ μ は反応強度の増幅を表す。Luce の定式化に従ってノード i の反応確率は次式:

$$p(R_i, t) = \frac{S_i(t)}{\sum_{j \in L} S_j(t)}, \tag{7}$$

ここで L は同一層内での競合するノードを表す。

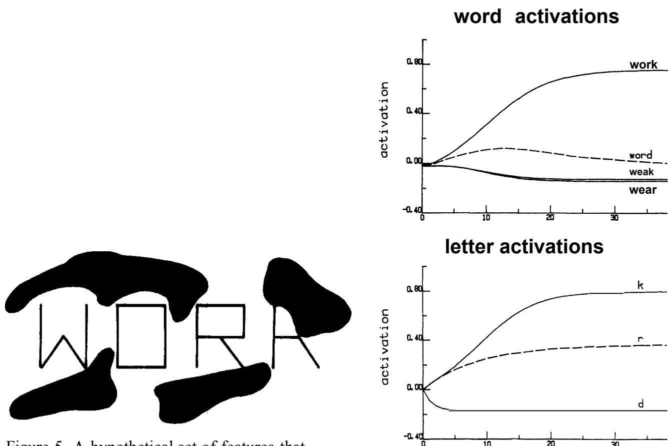


Figure 5. A hypothetical set of features that might be extracted on a trial in an experiment on word perception.

Figure 6. The time course of activations of selected nodes at the word and letter levels after extraction of the features shown in Figure 5.

図2　McClelland & Rumelhart (1981) より

McClelland & Rumelhart (1981) のシミュレーションでは 1179 実在単語が用いられた。刺激語は視覚特徴を表すベクトルへ変換する表を用いて入力ベクトルへと変換された。また空白は “-”，マスクは “0” で表された。以下のような刺激系列では，最初の 12 時刻まで mav- が提示され，12-23 時刻サイクルまで “mave” が提示され，最後に 24 時刻以降はマスクが提示されたことを意味する。

```
 0 mav-
12 mave
24 0000
```

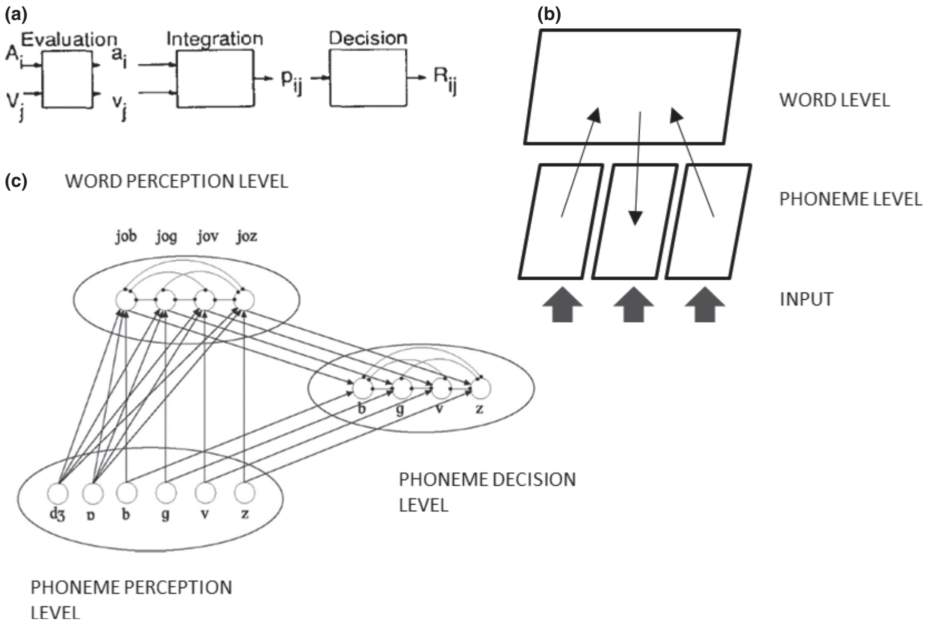


Fig. 8. (a) Massaro's schematic representation of the integration of stimulus and context information according to his Fuzzy Logical Model of Perception, reprinted from fig. 1, p. 401 of Massaro (1989). The A and V variables in Massaro's figure correspond to the stimulus and context variables presented in the text.(b) Schematic diagram indicating a unidirectional propagation of information for computing the contextual and stimulus factors used in Massaro's model for the identification of the segment in the middle position of a three-phoneme syllable. (c) The architecture of the MERGE model of speech perception (Norris et al., 2000), reprinted from fig. 11, p. 384 of Norris and McQueen (2008).

図3　From McClelland et al. (2014) Fig. 8

0.1 DIVA

CHAPTER

58

Neural Models of Motor Speech Control

Frank H. Guenther[1] and Gregory Hickok[2]

[1]Department of Speech, Language, and Hearing Sciences, Department of Biomedical Engineering, Boston University, Boston, MA, USA; [2]Department of Cognitive Sciences, Center for Language Science, Center for Cognitive Neuroscience, University of California Irvine, Irvine, CA, USA

58.1 INTRODUCTION

Speech production is a highly complex motor act involving the finely coordinated activation of approximately 100 muscles in the respiratory, laryngeal, and oral motor systems. To achieve this task, speakers utilize a large network of brain regions. This network includes regions involved in other motor tasks, such as the motor and somatosensory cortical areas, cerebellum, basal ganglia, and thalamus, as well as regions that are more specialized for speech and language, including inferior and middle prefrontal cortex and superior and middle temporal cortex. Our goal in this chapter is to describe the critical role of the auditory system in speech production. We first discuss the role of sensory systems in motor control broadly and summarize the long history of ideas and research on the interaction between auditory and motor systems for speech. We then describe current research on speech planning, which strongly implicates the auditory system in this process. Two large-scale neurocomputational models of speech production are then discussed. Finally, we highlight some future directions for research on speech production.

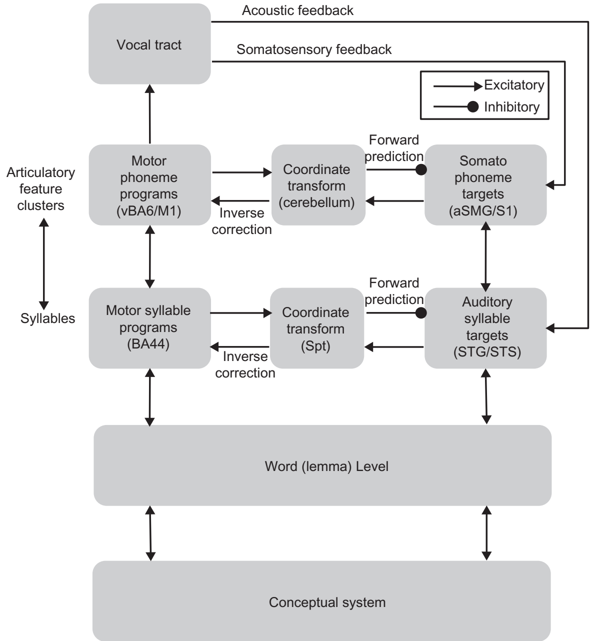
Movement is absolutely dependent on sensory information. We know where and how to reach for an object because we *see* its location and shape; we know how much force to exert while we are holding the object because we *feel* the pressure of the object on our hand and the weight on our limb; and we know how to initiate any of these movements because our sensory systems tell us where our limb is in relation to our body and the object. British neurologist and physiologist Henry Charlton Bastian (1837–1915) wrote on the topic of movement control in 1887, stating, "It may be regarded as a physiological axiom, that all purposive movements of animals are guided by sensations or by afferent impressions of some kind" (Bastian, 1887, p. 1). Experimental work over the decades backs these claims. This work has found, for example, that blocking somatosensory feedback from a monkey's limb (while leaving motor fibers intact) causes the limb to go dead. With training, the monkey can learn to reuse it clumsily, but only with visual feedback; blindfold the animal and motor control degrades dramatically (Sanes, Mauritz, Evarts, Dalakas, & Chu, 1984). Similar symptomology can be found in humans suffering from large-fiber sensory neuropathy, which deafferents the body sense while leaving motor fibers intact (Sanes et al., 1984).

Speech is no different. Without the auditory system, as in prelingual-onset peripheral deafness, normal speech development cannot occur. Importantly, it is not just during development that auditory information is critical. Experimental or naturally caused manipulations of acoustic input can have dramatic effects on speech production. For example, delayed auditory feedback induces nonfluency (Yates, 1963), altering feedback in the form of pitch or the formant frequency structure results in automatic and largely unconscious compensation in speech articulation (Burnett, Freedland, Larson, & Hain, 1998; Houde & Jordan, 1998; Larson, Burnett, Bauer, Kiran, & Hain, 2001), and exposure to a different linguistic environment can induce changes in the listener-speaker's articulation (picking up accents; Sancier & Fowler, 1997). Furthermore, although individuals who become deaf as adults can remain intelligible for years after they lose hearing, they show some speech output impairments immediately, including impaired ability to adjust pitch loudness in different listening conditions; over time, their phonetic contrasts become reduced (Perkell et al., 2000) and they exhibit articulatory decline (Waldstein, 1989).

The speech research literature contains numerous theoretical proposals that strongly link speech perception

Neurobiology of Language. DOI: http://dx.doi.org/10.1016/B978-0-12-407794-2.00058-4
725

0.2 HSFC



The HSFC model. The HSFC model includes two hierarchical levels of feedback control, each with its own internal and external sensory feedback loops. As in psycholinguistic models, the input to the HSFC model starts with the activation of a conceptual repre-sentation that, in turn, excites a corresponding word (lemma) representation. The word level projects in parallel to sensory and motor sides of the highest, fully cortical level of feedback control, the auditory–Spt–BA44 loop. This higher-level loop, in turn, projects, also in parallel, to the lower-level somatosensory–cerebellum–motor cortex loop. Direct connections between the word level and the lower-level circuit may also exist, although they are not depicted here. The HSFC model differs from the state feedback control (SFC) model in two main respects. First, "phonological" processing is distributed over two hierarchically organized levels, implicating a higher-level cortical auditory-motor cir-cuit and a lower-level somatosensory-motor circuit, which approximately map onto syllabic and phonemic levels of analysis, respectively. Second, a true efference copy signal is not a component of the model. Instead, the function served by an efference copy is integrated into the motor planning process. BA, Brodmann area; M1, primary motor cortex; S1, primary somatosensory area; aSMG, anterior supramarginal gyrus; STG, superior temporal gyrus; STS, superior temporal sulcus; vBA6, ventral BA6. The HSFC model is squarely within the tradition of the DIVA model in that it assumes that the targets of speech gestures are coded in auditory space and that feedback control is a key computa-tional operation of the network. HSFC differs from DIVA in three respects: (i) it assumes an internal as well as an external feedback detec-tion/correction mechanism; (ii) it situates auditory and somatosensory feedback loops in a hierarchical arrangement (auditory loop being higher-level and somatosensory loop being lower-level); and (iii) it assumes a modified computational architecture for the feedback loops.

0.3 Snodgrass

1 発話の計算モデル

1.1 SLAM

Walker & Hickok (2016) によれば SLAM とは semantic–lexical–auditory–motor model の省略形である。SLAM は前身モデル SP モデルFoygel & Dell (2000) に基づく。そして SP モデルは Dell (1986) を起源とする。論文 (Foygel & Dell, 2000) に従って Dell (1986) のモデルを DSMSG あるいは **重み=減衰 weigth-decay** モデルと表記する。 DSMSG (重み=減衰モデル) は失語症患者の絵画命名課題における誤答分布を説明するモデルである。DSMSG (重み=減衰モデル) は p と q の 2 つの大域パラメータを持つ。 p は拡散結合荷重であり，q は各ユニットの活性値の減衰を司るパラメータである。

Walker & Hickok (2016) によれば， Foygel & Dell (2000) に始まる一連の研究は 2 つのパラメータによって種々の変数が説明可能である。すなわち，絵画命名課題による臨床診断情報(Abel, Huber, & Dell, 2009)，語彙頻度効果(Kittredge, Dell, Verkuilen, & Schwartz, 2008)，失語症ごとに異なる文字誤答パタン(Schwartz, Dell, Martin, Gahl, & Sobel, 2006)，相互誤答パタン(Foygel & Dell, 2000) などが説明可能である。 2 つのパラメータとは意味=語彙間の重みである **s=重み s-weight** と，語彙=音韻表象間の重みである strong$p$=重み p-weight である。SP と命名されたアーキテクチャにより単語復唱課題 (Dell, Martin, & Schwartz, 2007)， 臨床画像に認められる神経学

的損傷部位からの予測 (Dell, Schwartz, Nozari, Faseyitan, & Branch Coslett, 2013) なども説明可能であるとされている。

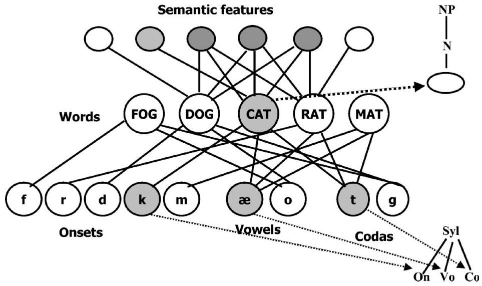


Fig. 1. Structure of the model.

図4　From Schwartz et al. (2006) Fig. 1, DSMSG, 重み=減衰モデル，あるいは SP モデルと呼ばれる

絵画命名検査のモデルであるが 4 に示すように絵画からの視覚入力は無視されている。もっぱら意味，単語，音韻の三者関係を取り扱っている。従って，Lichtheim2 や一連の読字モデルとの関連も指摘可能である。入力モダリティが意味から単語語彙表象を介して出力に至るというモデルであるから統合も可能であろう。

一方で，SLAM は，運動出力系との統合を試みている。 SLAM は患者のデータとして the Moss Aphasia Psycholinguistic Project Database www.mappd.org を用いている。

ところがおもちゃである。6 単語 (mat, rat, cat, dog, fog, log)，すなわち語彙ユニット数 6，各語彙層ニューロンは 10 個の意味層ニューロンと連結。意味的に類似単語間では 3 つの意味層ニューロンを共有する。意味層ニューロンの総数は 57。音韻層ユニット数は 10。その内訳は，オンセット 6，母音 2，コーダ 2。音韻的類似性を有する単語間ではオンセットのみが異なる。

SP モデルと SLAM の相違は，運動出力層が存在すること。語彙層から音韻層への結合に加えて，語彙層から運動層，音韻層から運動層への結合が存在する。

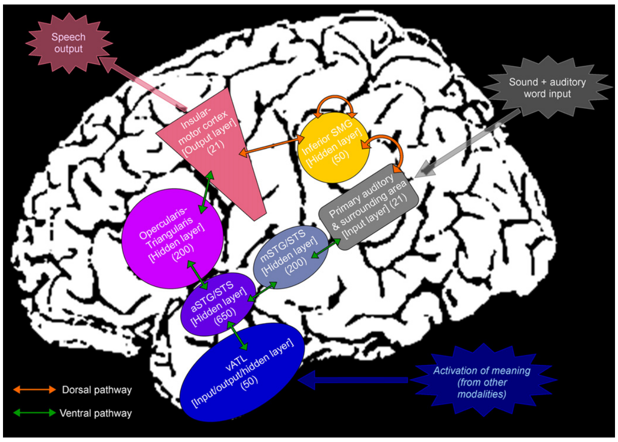
患者データは Moss Aphasia Psycholinguistic Project Data base www.mappd.org より。フィラデルフィア命名検査。

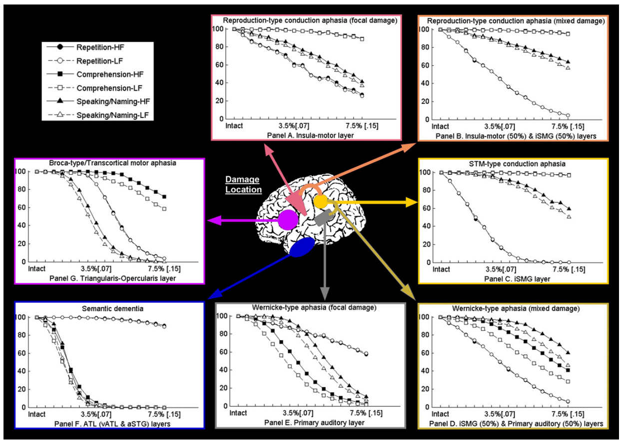
ユニット数と接続パターン: 6 つの単語のみで構成される 1 つがターゲットである非常に小さな語彙的近傍の構造を実装することにより英語の音声エラータイプの統計的確率を概算することを目的とた。モデルは 6 つの字句ユニットがあり，それぞれのユニットは意味的特徴を表す 10 個の意味ユニットに接続。意味的関連のある単語は 3 つの意味ユニットを共有している。つまり一般的な試行ではターゲットに意味的に関連する単語が 1 つだけでネットワークには合計 57 の意味ユニットがあります。各字句ユニットは、オンセット，母音，およびコーダに対応する 3 つの音韻単位にも接続。音韻ユニットは合計で 10 個あります。6 つのオンセット，2 つの母音，2 つのコーダ。ターゲットに音韻的に関連する単語は，そのオンセットユニットのみが異なり，ネットワークは常に 2 つのそのような単語で構成された。最後に，ネットワーク内の残りの 2 つの単語は、ターゲット、共有意味または音韻単位なし。試行の 20 % で音韻的に関連する 1 つの単語も意味的に関連し，ターゲットとの「混合」関係を持つ近隣を作成します。

解が命名シミュレーションは、意味ユニットに配信される活性値のブーストから始まる。 2 つのパラメーター S と P は字句の双方向結合を指定します。 2 つのパラメータとは字句-意味間結合と，字句-音韻間結合である。活性値の拡散は，全層へ同時に伝播する。活性値拡散は，ノイズと減衰を伴う線形活性則により 8 時刻，計算される。その後 2 番目の活性値ブーストが拡散される。語彙ユニットの活性値はさらに 8 時刻続く広がり続ける。最後に，最活性化した音韻オンセット，母音，コーダのユニットが，ターゲットと比較する出力として選択される。活性化レベルが低下するにつれてノイズの影響により生成エラーが発生する。これは強い結合によって軽減される。応答は，正しい，意味論的，形式的、混合、無関係，または新語として分類されます。与えられたパラメーター設定について、これらの6つの応答タイプにわたる多項分布は、モデルで多くの命名試行を生成することにより推定されます。これらの分布は、失語症患者によって生成されたネーミング応答から生じる分布と比較されます。

1.2 Lichtheim2

意味の表現が特殊。一つ前の正解を，腹側系の意味にエルマンネットの文脈層に入れることを持って意味と定義。





もちろん Lichtheim 2 は Lichtheim Lichtheim (1885) に基づく

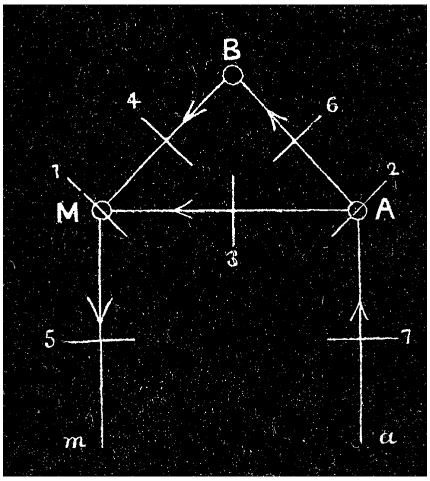


図5　Lichtheim (1885) Fig. 1 より

1.3 WEAVER

WEAVER: Word-form Encoding by Activation and VERification)

材料: CELEX ターゲット: 50 高頻度語, 妨害語: 50 CELEX 中からランダムに選んだ単語 (1) the forms of 50 nouns of highest frequency, and (2) the forms of 50 randomly selected nouns.

$$a\left(k, t+\Delta t\right) = a\left(k, t\right)\left(1-d\right) + \sum_{n} r a\left(n, t\right) \tag{8}$$

1.3.1 Logogen and Triangle models

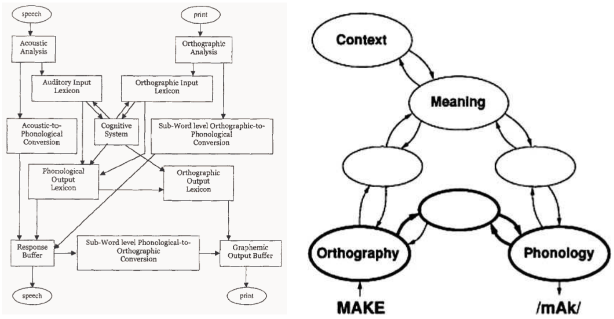
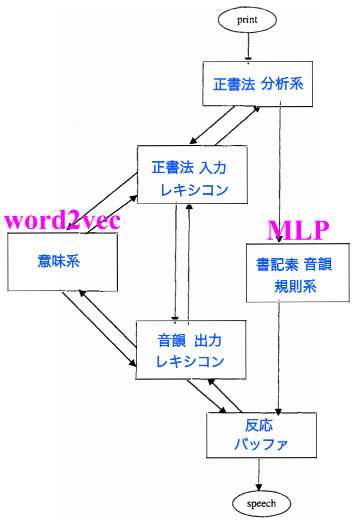


Figure 1. Two diagrams of the lexical processing system. The one on the left is taken from Patterson and Shewell (1987), and the one on the right is taken from Seidenberg and McClelland (1989). Lexical and sublexical pathways of processing can be found in both diagrams.

Triangle: Orhtography: 400 units, Phonology: 460, Hidden: 100/200, 2998 monosyllabic word

1.3.2 DRC



1.3.3 CDP++

刺激: (1) 32,270 形態-音韻単語対, (2) 8228 単音節語，24,042 双音節語

訓練データセット 30,519 形態-音韻単語対，7920 単音節語，22,596 双音節語,

http://ccnl.psy.unipd.it/CDP.html http://ccnl.psy.unipd.it/cdp_database

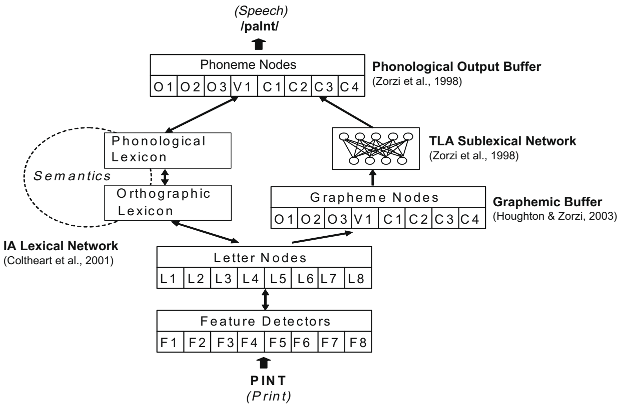


Fig. 2. The overall architecture of CDP+. Note: Numbers shown inside the various layers index slot positions, whereas letters indicate the type of representation (f = features, l = letter, o = onset, v = vowel, c = coda).

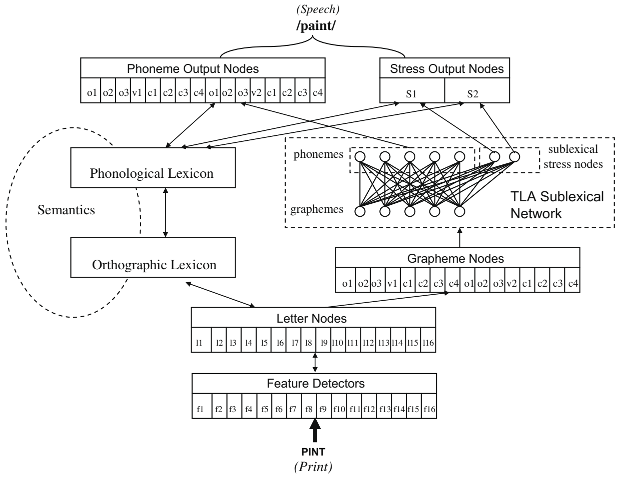


Fig. 3. The overall architecture of CDP++. Note: Numbers shown inside the various layers index slot positions, whereas letters indicate the type of representation (f = feature, l = letter, o = onset, v = vowel, c = coda). S1 = first syllable stress; S2 = second syllable stress.

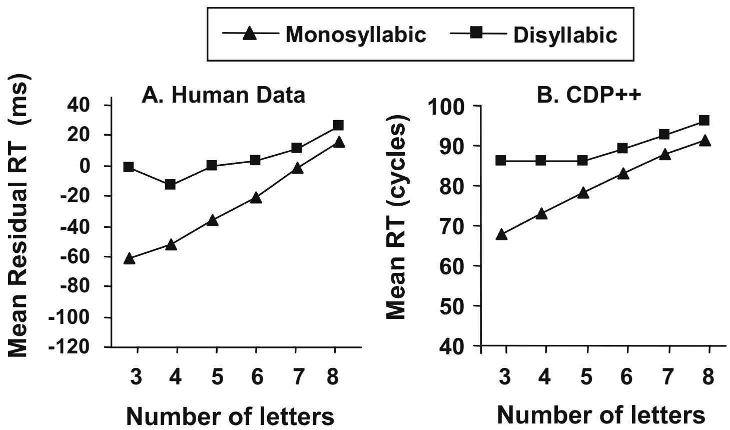


Fig. 4. Mean human and CDP++ reaction times (RTs) of monosyllabic and disyllabic words on the full ELP (2007) database.

Table 1
List of monosyllabic benchmark effects (from Perry et al. (2007)). Tick marks indicate successful simulations (for details, see Appendix D).

Name of effect	Description	CDP+	CDP++
Frequency	High-frequency words are faster/more accurate than low-frequency words	✓	✓
Lexicality	Words are faster/more accurate than pseudowords	✓	✓
Length × lexicality	Nonword naming latencies increase linearly with each additional letter	✓	✓
Frequency × regularity	Irregular words are slower/less accurate than regular words. This effect is typically larger for low-frequency words (Paap and Noel, 1991) but has also been reported for high-frequency words (Jared, 2002)	✓	✓
Word consistency	Inconsistent words are slower/less accurate than consistent words. The size of the effect depends on the friend–enemy ratio	✓	✓
Nonword consistency	Nonword pronunciations show graded consistency effects; that is, people do not always use the most common grapheme–phoneme correspondences	✓	✓
Position of irregularity	The size of the regularity effect is bigger for words with first position irregularities (e.g., *chef*) than for words with second or third position irregularities	✓	✓
Body neighborhood	Words with many body neighbors are faster/more accurate than words with few body neighbors	✓	–
Pseudohomophone advantage	Nonwords that sound like real words (e.g., *bloo*) are faster/more accurate than orthographic controls	✓	✓
Surface dyslexia	Patient MP showed a specific irregular word reading impairment that was modulated by the consistency ratio of the words as well as their frequency	✓	✓
Phonological dyslexia	Patient LB showed a specific nonword reading impairment which was reduced with pseudohomophones orthographically similar to their base words	✓	✓
Masked priming	Words preceded by a masked onset prime are faster/more accurate than words preceded by unrelated primes	✓	✓

Table 2
Percentage of variance accounted for (R^2) by CDP++, CDP+ (Perry et al., 2007), CDP (Zorzi et al., 1998a), the Triangle model (Plaut et al., 1996), and the DRC (Coltheart et al., 2001) on the Spieler and Balota (SB, 1997), Balota and Spieler (BS, 1998), Treiman et al. (1995), and Seidenberg and Waters (SW, 1989) databases.

Database	Models				
	CDP++	CDP+	CDP	Triangle	DRC
SB (1997)	19.5	17.3	5.9	3.3	3.7
BS (1998)	24.0	21.6	6.7	2.9	5.5
Treiman	18.1	15.9	6.5	3.3	4.8
SW	10.9	9.6	2.7	3.0	6.1

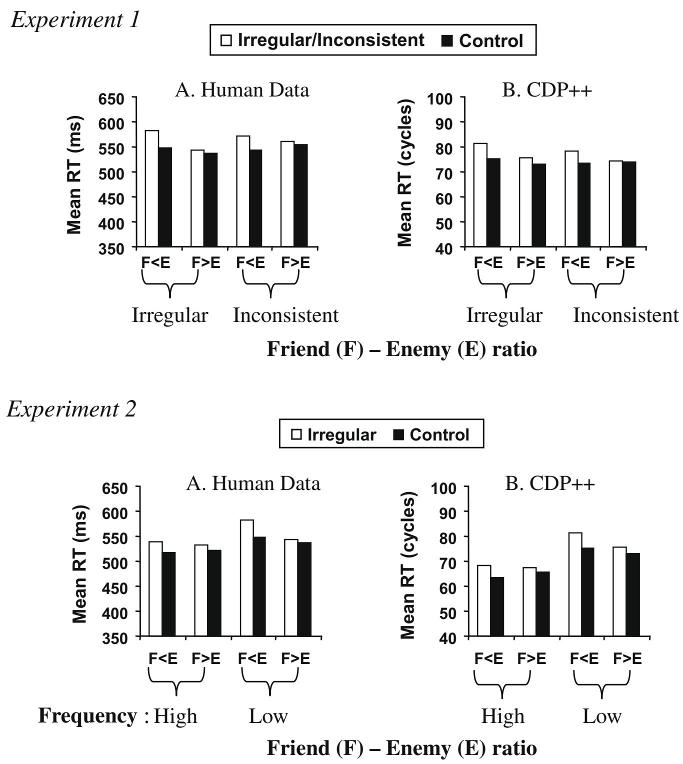


Fig. D1. Human data (milliseconds) and CDP++ simulations (cycles) of Jared's (2002) Experiment 1 and Experiment 2.

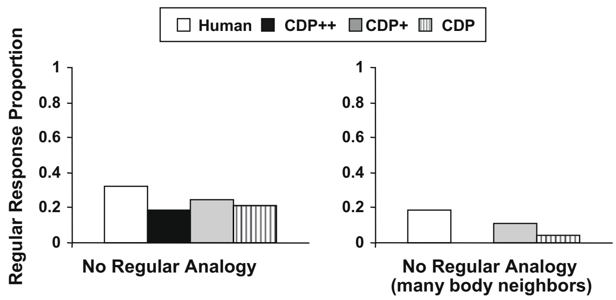


Fig. D2. Human data (response probabilities for regular pronunciations) and simulations of different models for the "no regular analogy nonwords" (Experiment 1) and the "no regular analogy with many body neighbors nonwords" (Experiment 2) of Andrews and Scarratt (1998).

1.3.4 Junction model

45,263 単語。 400 中間層ユニット。書記素 400，音韻 400，意味 400。

意味表現: 400 ビット。 LSA で作った。COALS データベース。Oxford handbook of Computational and Mathematical PsychologyNathan, Busemeyer, Wang, Townsend, & Eidels (2015) を参照。

1.3.5 Box 2 Semantic Memory Modeling Resources

A chapter on semantic models would seem incomplete without some code! Testing models of semantic memory has become much easier due to an increase in semantic modeling resources. There are now a wide variety of software packages that provide the ability to construct and test semantic models. The software packages vary in terms of their easy of installation and use, flexibility, and performance. In addition to the software packages, a limited number of Web-based resources exist for doing simple comparisons online. You may test models on standardized datasets, train them on your own corpora for semantic exploration, or use them for generating stimuli.

Software Packages

HiDEx (http://www.psych.ualberta.ca/westburylab/downloads/HiDEx.download.html): A C++ implementation of the HAL model; it is useful for constructing large word-by-word co-occurrence matrices and testing a wide variety of possible parameters.

SuperMatrix (http://semanticore.org/supermatrix/): A python implementation of a large number of semantic space model transformations (including PCA/SVD, Latent Dirichlet Allocation, and Random Vector Accumulation) on both word-by-word and word-by-document spaces. SuperMatrix was designed to emphasize the exchangeability of various sub-processes within semantic models (see Box 1), to allow isolation and testing the effects of specific model components.

GenSim (http://radimrehurek.com/gensim/): A python module that is very fast and efficient for constructing and testing word-by-document models, including LSA (reduced using SVD) and Topics (reduced using Latent Dirichlet Allocation).

S-Space (https://github.com/fozziethebeat/S-Space): A Java-based implementation of a large number of semantic space models, including HAL, LSA, BEAGLE, and COALS.

SEMMOD (http://mall.psy.ohio-state.edu/wiki/index.php/Semantic_Models_Package_(SEMMOD)): A python package to implement and compare many of the most common semantic models.

Word-Similarity (https://code.google.com/p/wordsimilarity/wiki/train): A tool to explore and visualize semantic spaces, displayed as directed graphical networks. Web-Based Resources

http://lsa.colorado.edu: The original LSA website provides the ability to exploreLatent Semantic Analysis with a wide variety of different metrics, including word-word similarities, similarities of passages of text to individual words, and similarities of passages of texts to each other.

http://semanticore.org: The Semanticore website is a web portal designed to bring data from many semantic models and psycholinguistic databases under one roof. Users can obtain frequency and co-occurrence statistics from a wide variety of corpora, as well as semantic similarities from a number of different semantic memory models, including HAL, LSA, BEAGLE, and Probabilistic Topics Models.

1.4 Sequence encoder model

訓練コーパス: 3751 単語。

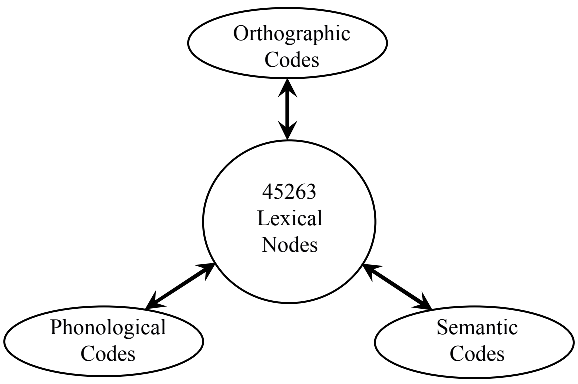


Figure 5. Basic architecture of the large-scale junction model.

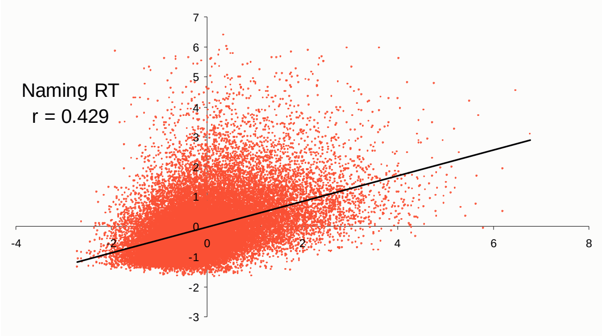


Figure 9. Model mean response times plotted against the naming response time residuals from the Elexicon database, in normalized coordinates.

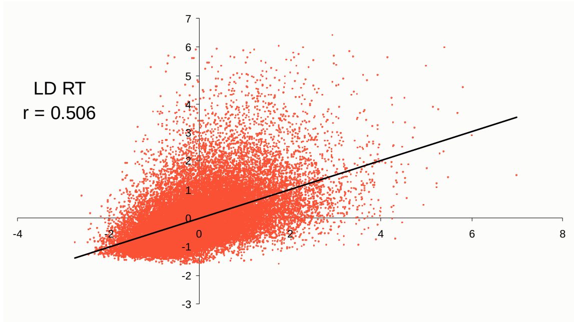


Figure 10. Model mean response times plotted against the lexical decision response time residuals from the Elexicon database, in normalized coordinates.

	DRC comparison		PMSP comparisons				
	N = 5190		N = 2808				
	Junction	DRC	Junction	Sim 1	Sim 2	Sim 3	Sim 4
R^2	**12.2%**	5.1%	**14.7%**	5.2%	4.1%	2.1%	11.9%

Table 1. Proportions of variance in naming response times accounted for by the junction model, compared with the DRC and PMSP models

1.5 Sequence encoder

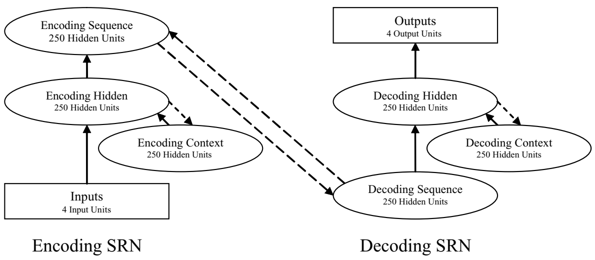


Fig. 1. Diagram of the sequence encoder architecture, with numbers of units used in Simulation 1 shown for each group. *Note:* These numbers were determined by trial and error to be sufficient to support near asymptotic performance on the training sequences. Solid arrows denote full connectivity and learned weights, and dashed arrows denote one-to-one copy connections. Rectangular groupings denote external (prescribed) representations coded over localist units, and oval groupings denote internal (learned) representations distributed over hidden units. SRN = simple recurrent network.

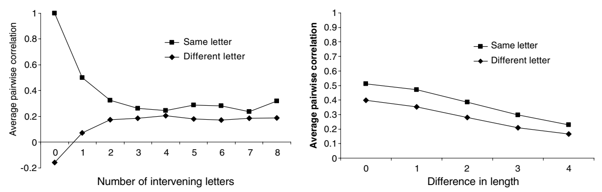


Fig. 2. Average pairwise correlations between conjunction patterns, plotted as a function of intervening letters (left) or difference in wordform length (right). *Note:* For intervening letters, the effect of wordform length was partialled out before correlations were computed.

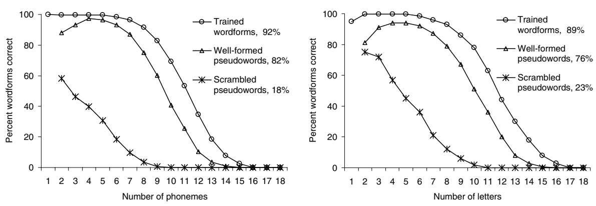


Fig. 3. Correct percentages for Simulations 2a and 2b, plotted as a function of wordform length and type.

- Dell (1986) SP のオリジナル, Dell, Schwartz, Martin, Saffran, & Gagnon (1997); Foygel & Dell (2000)
- WEAVER (Roelofs, A., 2014). A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. Cortex, 59, 33–48. doi:10.1016/j.cortex.2014.07.001
- cogsci, triangle, junction, sequnece encoder implement

1.6 cnps 勉強会 (2020-0128)のまとめ

- CDP++, ジャンクション(連接 junction), 系列符号化 (sequence encoder) モデルを簡単に概説
  - CDP++ モデルは DRC もでるの拡張版 (CDP+) の拡張版，書記素=音韻対応系を 2 層のニューラルネットワークモデルで置き換えて，単音節語のみならず多音節語，多言語，単語内の強調（ストレス）にも

対応したモデル
  - 系列符号化モデルは文字列の符号化と符号化した内部表現を復号化（音韻系列）する 2 つの単純再帰型ニューラルネットワークモデルをつなぎ合わせたモデル
  - 連接モデルは，書記素，音韻，意味の 3 つの構成モジュールを中央の語彙(レキシコン) に連接（ジャンクション）するモデル。三角モデルが，各構成モジュールを個別につなぐモデルであるのに対して，中央の語彙表象を仮定する点が異なる。
  - 三角モデルと連接モデルとをあわせるとテトラポッド（四角錐）モデルになるのではないだろう（というのは浅川の妄想）
- 2018 年は BERT モデルが自然言語処理業界でパンデミック。グーグルの検索エンジンにも組み込まれて検索精度が向上した。日本語化した BERT も存在するので後日紹介します。
  - BERT 以外に ELMO があって，どちらもセサミストリートのキャラクターから命名されました。
-

ついでに SLAM とか DSMSG とかがあるのであった。

引用文献

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204-256.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283-321.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, *43*, 182–216.

Fujimura, O. (1992). Phonology and phonetics - a syllable-based model of articulatory organization. *Journal of Acoustic Society Japan*, *13*, 39-48.

Fujimura, O. (2000). The c/d model and prosodic control of articulatory behavior. *Phonetica*, *57*, 128-138.

Fujimura, O. (2003). The c/d model: A progress report. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of 15th International Congress of Phonetic Sciences (ICPhS-15)* (p. 993-996). Barcelona, Spain.

Guenther, F. H. (2016). *Neural control of speech.* MIT Press.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145.

Hickok, G., & Small, S. L. (Eds.). (2016). *Neurobiology of language.* Elseiver.

Kello, C. T. (2006). Considering the junction model of lexical processing. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 50–75). Psychology Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*, 1-38.

Lichtheim, L. (1885). On aphasia. *Brain*, 433–484.

McClelland, J. L., Mirman, D., Bolger, D. J., & Khaitan, P. (2014). Interactive activation and mutual constraint satisfaction in perception and cognition. *Cognitive Science*, *38*, 1139-1189.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception, i: An account of basic findings. *Psychological Review*, *88*, 375–407.

Nathan, P. E., Busemeyer, J. R., Wang, Z., Townsend, J. T., & Eidels, A. (Eds.). (2015). *The oxford handbook of computational and mathematical psychology.* Oxford University Press.

Norris, D. (2013). Models of visual word recognition. *Trends in Cognitive Sciences*, *17*, 517–524.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology*, *61*, 106–151.

Plaut, D., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

Roelofs, A. (1997). The weaver model of word-form encoding in speech production. *Cognition*, *64*, 249—284.

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: The weaver++/arc model. *Cortex*, *59*, 33-48.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. the contextual enhancement effect and some tests and extension of the model. *Psychological Review*, *89*, 60–94.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*, 228–264.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmenta model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Sibley, D. E., Kello, C. T., Plaut, D. C., & Elman, J. L. (2008). Large-scale modeling of wordform learning and representation. *Cognitive Science*, *32*, 741–754.

Tourville, J. A., & Guenther, F. H. (2011). The diva model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*, 952-981.

Walker, G. M., & Hickok, G. (2016). Bridging computational approaches to speech production: The semantic–lexical–auditory–motor model (SLAM). *Psychonomic Bulletin and Review*, *23*, 339-352.

Walker, G. M., Hickok, G., & Fridriksson, J. (2018). A cognitive psychometric model for assessment of picture naming abilities in aphasia. *Psychological Assessment*, *30*, 809-826.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? a connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1131–1161.