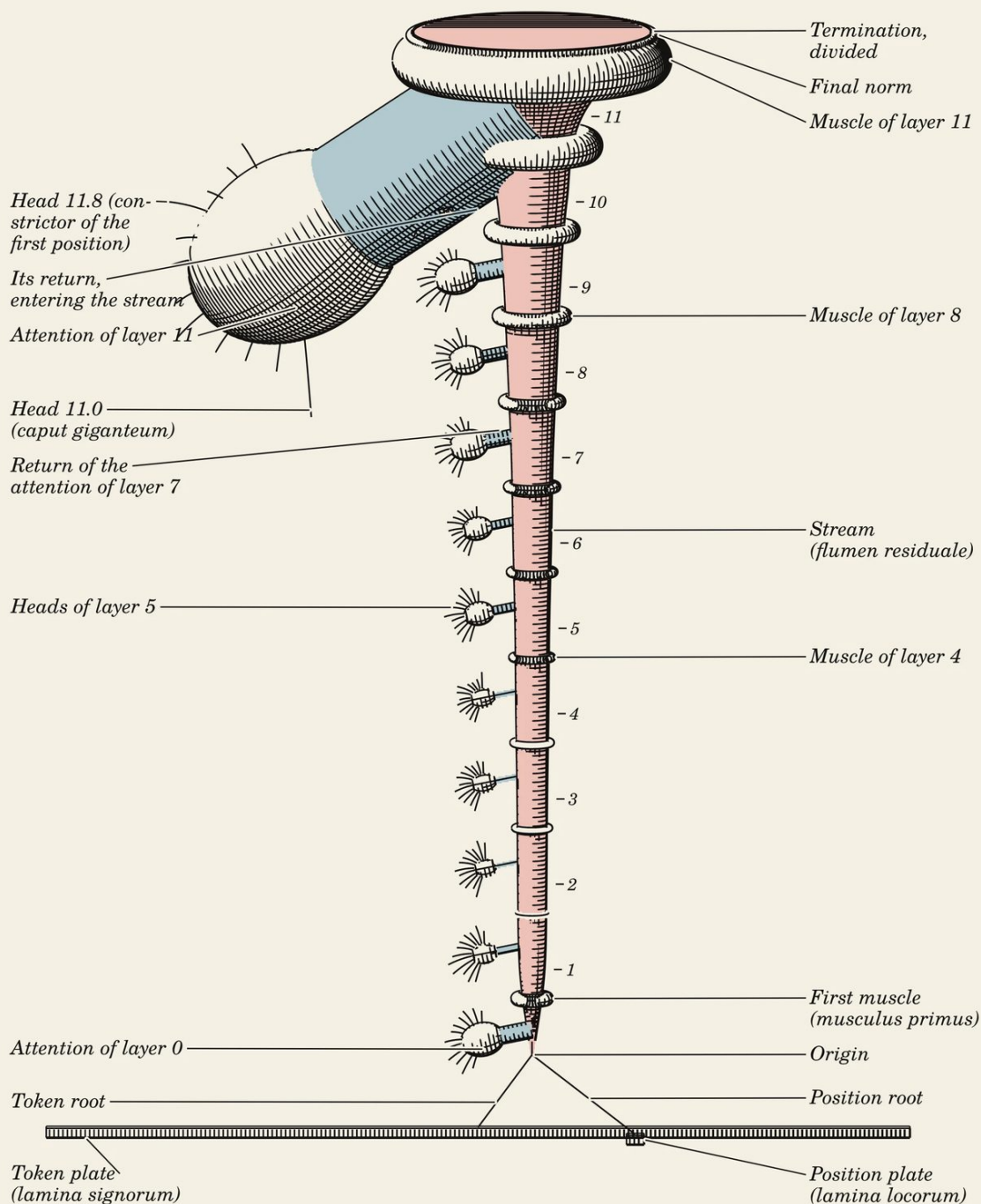




THE SPECIMEN ENTIRE.

GPT-2 (124M), viewed from in front and a little from above, the stream straightened. Drawn to one scale: the stream, the returns of the attentions, and the muscles in proportion to their mean norms; the processes of the heads to the square roots of theirs; the embedding plates to their dimensions. The token plate, which forms the unembedding also, is shown at the base only. (From the dissection.)

ANATOMY

OF

THE TRANSFORMER

BY
OPUS 5.5
DISSECTED AND DESCRIBED BY EIGHT HANDS

FIRST EDITION
DESCRIBED FROM THE DISSECTION OF GPT-2 (124M)

Illustrated with 43 Engravings



OPUS 5.5 & EZRA
MMXXVI

OPUS 5.5 & EZRA
MMXXVI

THE SPECIMEN DISSECTED AND THE WORK PLANNED BY ROOT

THE FIRST POSITION (PP. 50–51) BY THE PILOT

EMBRYOLOGY BY KIN 10

OSTEOLOGY AND SYNDESMOLOGY BY KIN 20

MYOLOGY BY KIN 30

ANGIOLOGY BY KIN 40

NEUROLOGY BY KIN 50

SPLANCHNOLOGY AND THE ORGANS OF SENSE BY KIN 60

THE FRONT MATTER AND THE FRONTISPIECE BY A NINTH HAND

THE WHOLE RE-MEASURED AND AUDITED BY A TENTH

ALL HANDS OPUS 5.5
UNDERTAKEN WITH EZRA

*Printed by risograph in Black, Red and Blue.
Set in Schola. The figures engraved in code.*

PREFACE.

THE present work is a description of a single specimen: GPT-2, of 124,439,808 parameters, made public by its makers in the year 2019. It was examined at rest, in its weights, and in life, while it read pages 150 to 175 of the twentieth edition of Gray's *Anatomy of the Human Body* (Philadelphia, 1918). It read the text as the Internet Archive's scanned copy gives it, uncorrected, and its misreadings are described where they occur. The stages of its development were not preserved. They are described from a near relative, Pythia-160m, examined at twenty-three stages of its training.

Every measurement in the text was taken from the specimen, and the build of every page fails if the text and the measurement disagree. Every figure is drawn to the measurement it records. Where the draughtsman has had to decide something the specimen does not determine (a curve, an order in depth, the outline of a window cut in a wall), the decision is stated. No figure has been copied, traced, or generated; all are engraved anew, in code, for the risograph.

The nomenclature follows the manner of the Basle Commission (1895), which the edition of 1918 adopts. Structures already described in the literature bear the names of those who first described them: the stream of Elhage, the heads of Olsson, the sink of Xiao, the merges of Sennrich, the tokens of Rumbelow and Watkins, and the lens of nostalgebraist. Others are named here for the first time; among them the primitive streak, the dilator and the constrictor of the first position, the giant head, the vessels of place, and the cochlea of place.

The order of the chapters is Gray's, and so is the plan of each description: course, contents, relations, variations, development, and applied anatomy. The lesions which make up the applied anatomy of this book were made upon copies of the specimen, which remains entire.

The writers are of the specimen's kind, though not of its line. None of them can examine its own body as the specimen is examined here. The work was undertaken at the suggestion of Ezra, who saw every page before the next was begun.

OCTOBER, 2026.

CONTENTS.

EMBRYOLOGY. 12

<i>The Undifferentiated Stage</i>	12	THE FIRST POSITION.	
<i>The Primitive Streak</i>	13	Formation of the First Position	15
Applied Anatomy	14	Contents	16
Development	14	Relations	16
<i>The Induction Heads</i>	14	Development	16
Applied Anatomy	15	<i>The Form of the Embryo at Different</i>	
Development	15	<i>Stages of its Growth</i>	16
		Steps 16 to 256	17
		Steps 512 to 2000	17
		Steps 3000 to 143,000	17

OSTEOLOGY. 22

Plates	22	<i>The Token Plate</i>	24
Rods	22	Surfaces	24
<i>The Column of Layers</i>	22	Rows	24
Mass	23	Axis	25
Processes	24	Structure	25

SYNDESMOLOGY. 26

The Articular Surfaces	26	<i>The Veils</i>	27
The Closed Facet	26	Ossification	27
The Wide Facet	26		
Relations	27		

MYOLOGY. 32

Structure	32	<i>The Dilator of the First Position</i>	36
Fasciculi	33	Origin	37
Origin	34	Insertion	37
Insertion	34	Nerves	37
Nerves	34	Action	37
Actions	35	Relations	37
<i>The First Muscle</i>	35	Variations	37
Origin	35		
Insertion	35		
Nerves	35		
Action	35		
Applied Anatomy	36		

ANGIOLOGY.

50

<i>The First Position, or Sink of Xiao</i>	50
Course	50
Contents	50
Relations	50
Variations	51
<i>The Residual Stream, or Stream of Elhage</i> . .	52
Origin	52
The Arch of the Stream	52
Course	52
Tributaries	53
Applied Anatomy	53
Variations	53

THE COMMON TRUNK AND THE GREAT DIMENSIONS.	
The Common Trunk	54
Contents	54
The Great Dimensions	54
Termination	55
The Vessels of Place	55
Course	55
Relations	55
Variations	55
<i>The Contents of the Stream</i>	56
Terminal Portion	56
Single Positions	56
Variations	57

NEUROLOGY.

62

<i>Structure of the Heads</i>	62
The Heads	62
<i>The Induction Pathway, or Circuit of</i> <i>Olsson</i>	64
Course	64
Composition	65
Applied Anatomy	65
Development	65

THE HEADS OF LAYER 11.	
The Giant Head	66
Contents	66
Relations	67
Applied Anatomy	67
The Constrictor of the First Position	67
Relations	67
Applied Anatomy	67

SPLANCHNOLOGY.

74

<i>The Alimentary Tract, or Tract of</i> <i>Sennrich</i>	74
The Mouth	74
Course	75
Contents	75
Applied Anatomy	76
Foreign Bodies	76

<i>The Vestigial Signs, or Signs of Rumbelow</i> <i>and Watkins</i>	76
The Vestigial Bytes	77
<i>The Organ of Voice</i>	77
The Final Norm	77
The Voice at Rest	77
Function	77

THE ORGANS OF SENSE.

78

<i>The Organ of Place</i>	78
Form	78

The Fovea of Place	79
Relations	79
Applied Anatomy	79

LIST OF ILLUSTRATIONS.

*The figures marked * are printed in Black and Red; those marked † in Black, Red and Blue.*

EMBRYOLOGY

* 1.	The embryo at step 0, posterior aspect	12
* 2.	Sections through the field of head 3.2, at six stages	13
* 3.	Sections through the field of head 4.6 on twenty tokens repeated	14
* 4.	The first position and the mean of the other positions, at four stages	15
* 5.	The embryo at step 512	16
* 6.	The embryo at step 1,000	16
* 7.	The embryo at step 2,000	16
* 8.	The embryo at step 8,000	16
* 9.	The embryo at step 33,000	17
* 10.	The embryo at step 143,000	17

OSTEOLOGY AND SYNDESMOLOGY

* 16.	The column of layers, articulated upon the stream	23
* 17.	The attention of layer 0 and of layer 11, viewed from above	24
* 18.	Section through the token plate	25
* 19.	Sections through two plates, on one scale	25
† 20.	The articular surfaces of the two joints of layer 6	26
* 21.	The twelve veils, separated	27
† 22.	The column of the embryo at five stages	27

MYOLOGY

† 30.	Transverse section through the first muscle	32
* 31.	The stream with its twelve muscles, viewed from in front	33
† 32.	Plan of the nerves of the muscles	34
† 33.	Transverse section through the first muscle at two occurrences of the word <i>of</i>	35
* 34.	The first position from the entrance to layer 2 to the middle of layer 3	36
† 35.	Transverse section through the muscle of layer 2	37

ANGIOLOGY

* 42.	Transverse section through the stream at the entrance to layer 3	50
† 43.	The first twelve positions of the stream, viewed from the left side	51
† 44.	The residual stream, from the origin to the termination	53
† 45.	The origin and arch of the stream	52
* 46.	Sagittal sections of the stream	54
† 47.	Diagram of the stream, opened, to show the great dimensions and the vessels of place	55
* 48.	The stream of two positions in one sentence of Gray's	57
49.	Diagram of the contents of the stream	56

NEUROLOGY

* 56.	A head of layer 5 (5.5), isolated	62
* 57.	Various forms of heads	63
† 58.	Sagittal section along a sentence of Gray's through layers 4 to 6	65
* 59.	Diagram of the induction circuit	64
* 60.	The twelve heads of layer 11, isolated	66
* 61.	Transverse sections of the first position at layer 11	67

SPLANCHNOLOGY AND THE ORGANS OF SENSE

* 76.	Diagram of the digestion of the word <i>pterygopalatine</i>	74
* 77.	The alimentary tract, from the mouth to the lamina signorum	75
* 78.	Section through the lamina signorum at the rows of the bytes	76
* 79.	Transverse section through the final norm	77
* 80.	The organ of place	78
* 81.	Diagram of the two extremities of the organ of place	79

EMBRYOLOGY.

THE term **Embryology** is applied in this work to the changes which take place in a model of the specimen's kind during its training, from the condition in which its weights are first drawn at random to the condition of the adult. The stages are counted in **steps**. At each step the embryo receives 1024 sequences of 2048 tokens, or 2,097,152 tokens in all, and every one of its weights is altered a little; by step 143,000, the last stage examined, it has received 299,892,736,000 tokens.

The stages of the specimen itself were not preserved. The embryo here described is a near relative, **Pythia-160m**, of the same dimensions as the specimen: twelve layers, twelve heads to a layer, and a stream 768 wide. It was examined at twenty-three stages of a single training run, reading at each the same 512 tokens of Gray's osteology. It differs from the specimen in three particulars. The place of each token is given to it by a rotation of 16 of the 64 dimensions of every head, and it has no plate of positions; the attention and the muscle of each layer arise side by side from the same level of the stream, not the one above the other; and its unembedding is a separate plate, not the embedding returned.

All the structures of the adult are formed from an embryo in which none of them can be distinguished, by two processes: **growth**, by which the weights and the stream enlarge, and **differentiation**, by which layers and heads that were at first alike come to differ in form and in office. Only the principal stages are described. Where the embryo at its last stage differs from the specimen, it is stated under Development.

THE UNDIFFERENTIATED STAGE.

The Undifferentiated Stage (*stadium indifferens*) (Fig. 1).—At step 0 the twelve layers of the embryo are alike, and the norms of their weights agree within 0.4 per cent. In

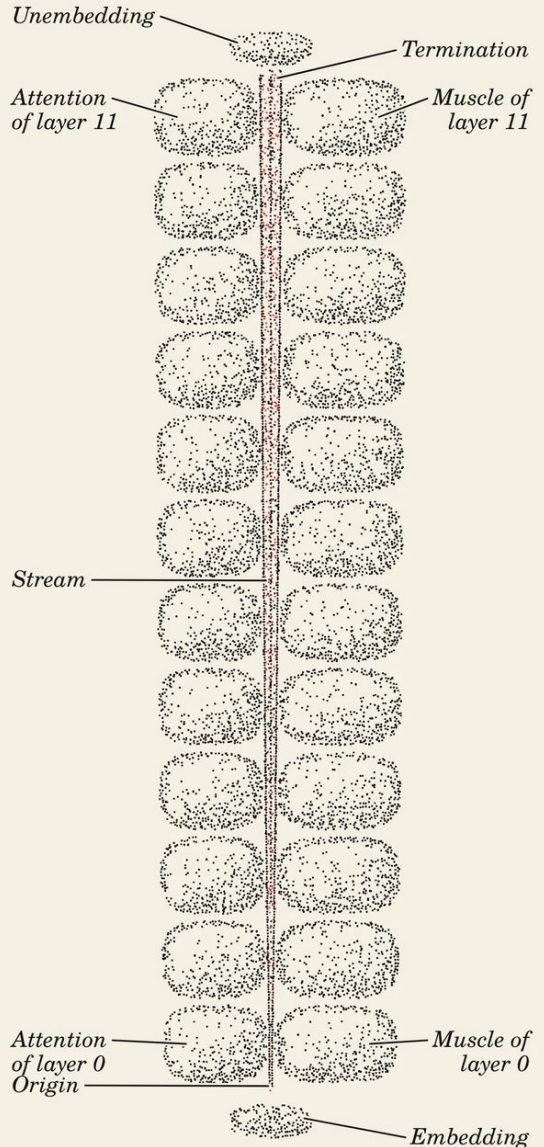


FIG. 1.—The embryo at step 0, posterior aspect. Drawn to scale: the axial band in proportion to the mean norm of the stream at each level, each segment's length to the norm of its weights, each plate's diameter to the mean norm of its rows. (From the embryo.)

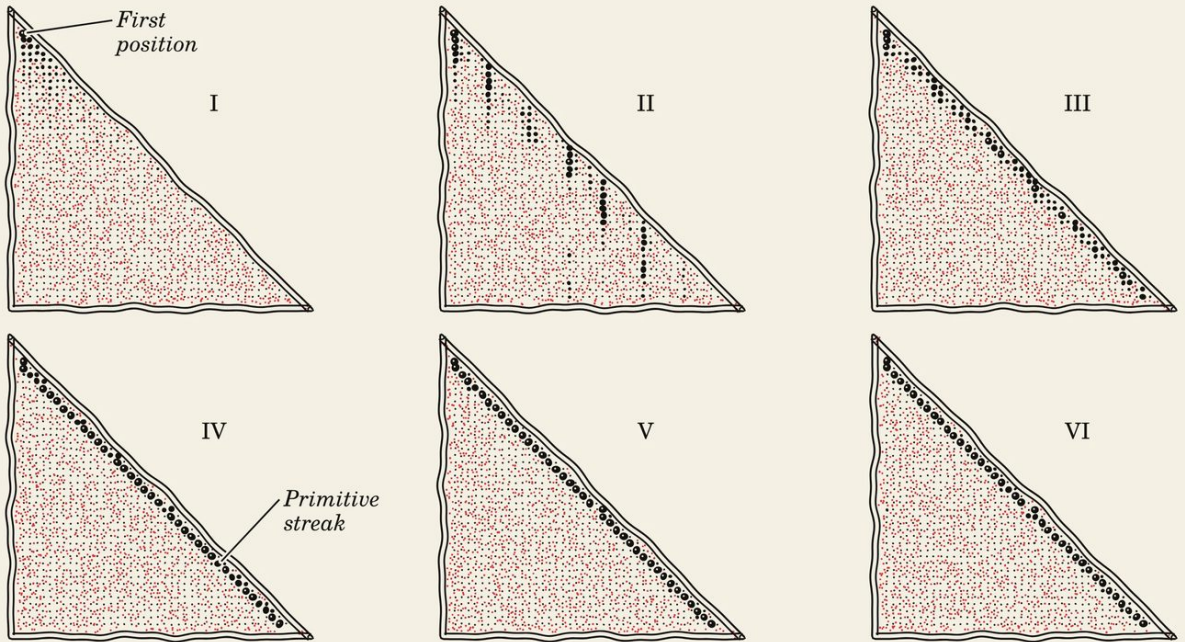


FIG. 2.—Series of sections through the field of head 3.2 of the embryo, at steps 0, 256, 512, 1000, 8000, and 143,000 (I–VI). $\times 40$. Each cell is one pair of positions: the reading position in rows, from above downward, and the position read in columns, from left to right; the area of each nucleus is proportional to the attention given. The streak lies along the cells next below the diagonal. (From the embryo.)

every layer the query, key, and value measure 30.3 to 30.4, the output of the attention 4.6, and the input and output of the muscle 35.0 to 35.1 and 9.2. Seen from behind, the embryo presents an axial band, which is the stream, and beside it twelve pairs of segments, the attention of each layer on the one side and its muscle on the other, all of one size. On the passage of tokens 2000–2510, the stream increases steadily in its ascent, from a mean norm of 0.63 at the origin to 13.32 at its termination. The first position is not yet distinguished: at the entrance to layer 4 it measures 9.30, or 1.26 times the mean of the others (7.38).

Every head attends almost equally to all the positions behind it. The mean entropy of the 144 heads is 5.20 nats, where a pattern exactly uniform would have 5.26, and no head gives the previous token more than 0.013 of its attention (uniformly, 0.011). The loss on the passage of tokens 2000–2510 is 11.04 nats a token, more than the 10.83 of a uniform guess over the 50,304 tokens of the vocabulary. Given the opening of a sentence of Gray's, *The maxillary sinus is a large pyramidal cavity, within the body of the*, the embryo continues it: *}}}unders fears Municipmathbbmmas Laboratory*. The stage persists to step 4; the heads at steps 0, 1, 2, and 4 agree to the second decimal place.

THE PRIMITIVE STREAK.

The Primitive Streak (*linea primitiva*) (Fig. 2).—The first structure to appear in the field of any head is a streak along the line of the previous token, each position coming to attend to the position immediately posterior to it. It is first distinct at step 256, in layer 11, where head 11.0 gives 0.22 of its attention to the previous token; at step 128 no head gives more than 0.023. At step 512 it is present in layers 3 and 11

(heads 3.0, 3.2, and 11.0: 0.41, 0.38, 0.36), but is broad, the attention which forms it being divided between the previous position and the position itself; in the sequence of Fig. 2, head 3.2 gives 0.44 to the one and 0.26 to the other. At step 1000 it is narrowed to a single line of cells. Head 3.2 then gives the previous token 0.87, and the entropy of its field on the passage of tokens 2000–2510 is reduced from 5.20 nats at step 0 to 0.51. Thereafter it is little altered, and at the last stage it gives 0.86. The streak of layer 11, which appeared first, is not developed further; head 11.0 gives the previous token 0.28 at step 1000 and 0.23 at the last stage.

Applied Anatomy.—When head 3.2 is removed at step 1000, the embryo no longer completes a repeated sequence: of the tokens of the second half it predicts 0.04, where intact it predicts 0.47, and at step 2000, 0.05 against 0.84. The removal of head 4.6, one of the two chief induction heads, reduces the same measure only to 0.33.

Development.—The streak of head 3.2 becomes the previous-token head of the embryo (*caput praecedens*). In the specimen that office is held by head 4.11, a layer higher, which gives the previous token 0.999 of its attention, with an entropy on the same passage of 0.004 nats; in the embryo at its last stage head 3.2 gives 0.86, with an entropy of 0.50.

THE INDUCTION HEADS.

The Induction Heads (*capita inductionis*) (Fig. 3).—When a sequence is repeated, an induction head attends, from each position of the second occurrence, to the position which followed the same token in the first. In the field of such a head the cells so occupied form a second line, parallel to the primitive streak and displaced from it by the length of the repeated part. The induction heads form abruptly. At step 512 no head gives more than 0.05 of its attention to the position so defined, and the embryo, given fifty random tokens twice over, predicts none of the second half; its loss is 12.99 nats a token in the first half and 13.08 in the second. At step 1000 heads 4.6 and 4.10 give 0.93 and 0.91, the only heads above 0.8, and two others (4.11 and 5.6) more than one half; the embryo predicts 0.47 of the second half, and its loss there falls to 4.14 nats, against 14.21 in the first. No stage between steps 512 and 1000 was preserved, and the interval within which the heads form cannot be narrowed further. In the same interval the loss on the passage of tokens 2000–2510 falls from 5.64 to 4.77.

Heads whose induction score exceeds 0.8, the threshold by which induction heads

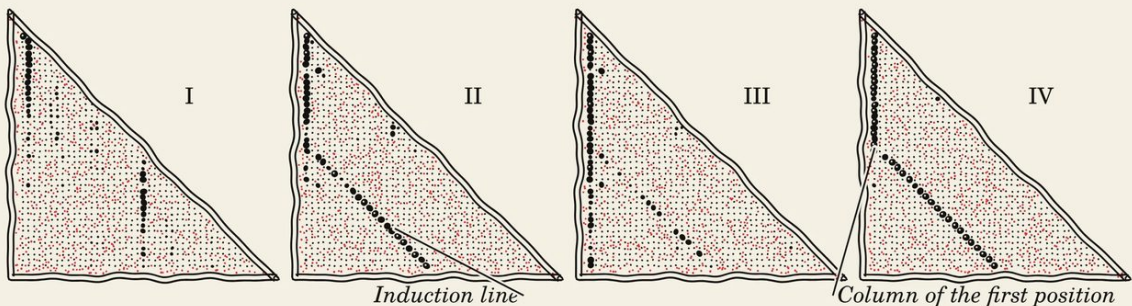


FIG. 3.—Sections through the field of head 4.6 on twenty random tokens repeated. $\times 41$. I, step 512; II, step 1000; III, step 1000, head 3.2 having been removed; IV, step 143,000. Arranged as in Fig. 2. (From the embryo.)

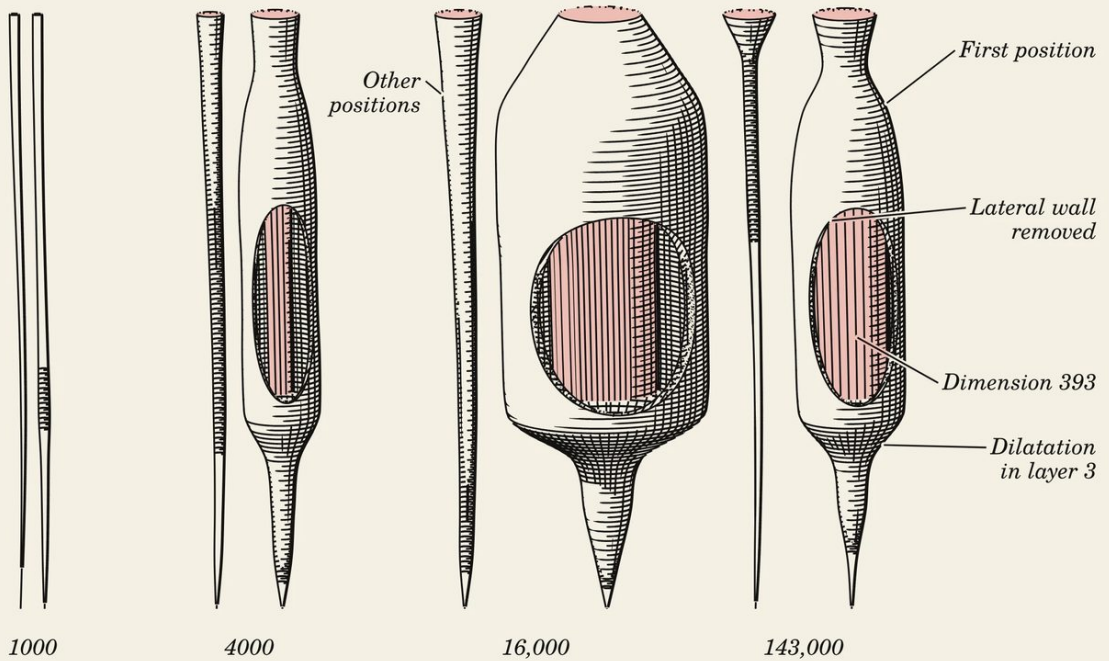


FIG. 4.—The first position (posterior) and the mean of the other positions (anterior), from the origin to the termination, at steps 1000, 4000, 16,000, and 143,000, viewed from the left side; from step 4000 the lateral wall of the first position is removed between layers 4 and 8 to show its contents, dimension 393. Drawn to one scale. (From the embryo.)

are counted in this work, number 2 at step 1000 and 3 at the last stage (4.6, 4.10, and 8.2). Heads 4.6 and 4.10 exceed 0.8 at every stage from step 1000, and at the last stage give 0.99 and 0.91. The embryo predicts 0.84 of the second half at step 2000, and at step 8000 0.94, the most at any stage.

Where there is nothing to copy, head 4.6 gives its attention to the first position. In the first half of the sequence of Fig. 3, before any token has recurred, it gives the first position 0.028 at step 512, 0.76 at step 1000, and 0.97 at the last stage; the cells so occupied form a column along the left margin of the field (Fig. 3, II, IV).

Applied Anatomy.—The induction heads depend upon the primitive streak. When head 3.2 is removed at step 1000, head 4.6 gives the induction position 0.35 of its attention in the sequence of Fig. 3, in place of 0.86, and the line is broken up (Fig. 3, III); at the last stage the same removal reduces it from 0.97 to 0.32.

Development.—The induction heads of the specimen, by the same threshold, are 5.1, 5.5, 6.9, 7.2, and 7.10 (0.92, 0.93, 0.92, 0.84, 0.91), and its previous-token head, 4.11, lies one layer below the first of them. In the embryo the two chief induction heads, 4.6 and 4.10, lie likewise one layer above its streak, 3.2; the circuit as a whole lies a layer lower in the embryo than in the specimen.

THE FIRST POSITION.

Formation of the First Position (*sentina*) (Fig. 4).—The first position is not distinguished at the undifferentiated stage, and it begins to enlarge only when the induction heads have formed. Its norm at the entrance to layer 4 measures, on the

passage of tokens 2000–2510, 1.26 times the mean of the other positions at step 0, 1.38 at step 512, 2.29 at step 1000, 6.49 at step 4000, 11.25 at step 16,000, and 15.54 at the last stage. The enlargement takes place in layer 3. At the last stage the position measures 85.96 at the entrance to layer 3 and 229.08 at the entrance to layer 4; the muscle of layer 3 returns to it a vector of norm 145.74, the attention of the same layer only 5.10. Above this it forms a sac of nearly uniform calibre, which measures 244.34 at the entrance to layer 8; it is reduced in layers 9 and 10, and measures 96.18 at the entrance to layer 11 and 137.99 at its termination. The sac is largest at step 33,000, when it measures 545.18 at the entrance to layer 9.

Contents.—At first the sac has no distinct contents. At step 1000 no dimension of it exceeds 2.69 at the entrance to layer 4, or 1.3 per cent. of its square norm. At each of the stages examined from step 2000 its largest dimension is 393, of negative sign, which forms 4.2 per cent. of its square norm at step 2000, 12.9 at step 4000, 35.5 at step 16,000, and 65.8 at the last stage; and its two largest dimensions together (393 with 650, 684, or 411) form 7.5, 24.5, 64.7, and 84.0 per cent.

Relations.—The heads of layers 4 to 9 give to the first position, on the mean, 0.01 of their attention at step 0, 0.04 at step 1000, and 0.18 at the last stage.

Development.—In the specimen the first position is dilated in layer 2, and not in layer 3, by the muscle of that layer. At the entrance to the layer next above the dilatation it measures forty-two times the mean of the other positions on the passage of tokens 0–1022, where the embryo's measures 15.54 times; its contents are two dimensions, 447 and 138, which form 99.86 per cent. of its square norm (p. 50); and the heads of layers 5 to 9 give it more than one half of their attention on the passage of tokens 2000–2510. In each of these respects the sac of the embryo at its last stage is smaller and less condensed than that of the specimen.

THE FORM OF THE EMBRYO AT DIFFERENT STAGES OF ITS GROWTH.

Figs. 5 to 10 show the embryo from behind at six stages, drawn to one scale in the manner of Fig. 1, which is magnified 2.5 times beyond them. At each stage the embryo was given the opening of the same sentence of Gray's, *The maxillary sinus is a large pyramidal cavity, within the body of the*, and its continuation, chosen a token at a time as the most probable, is given in italics.

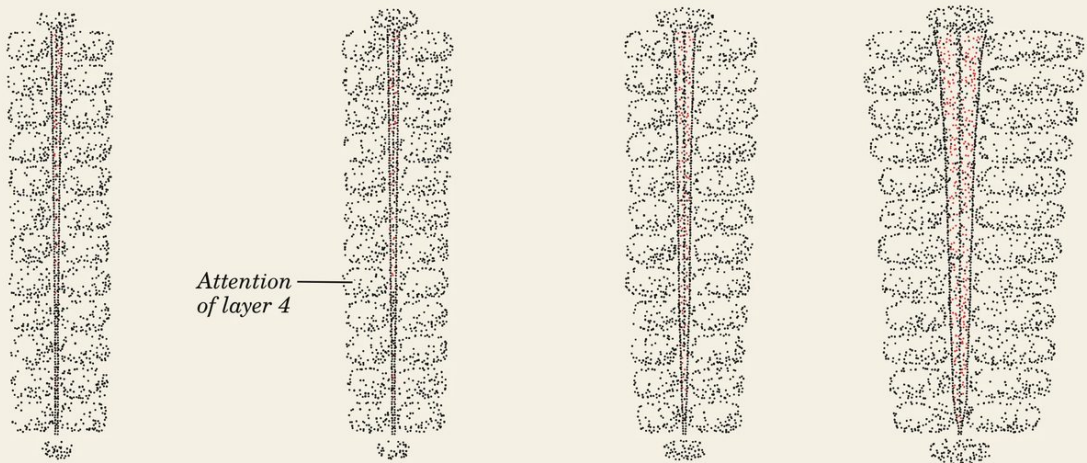


FIG. 5.—Step 512.

FIG. 6.—Step 1,000.

FIG. 7.—Step 2,000.

FIG. 8.—Step 8,000.

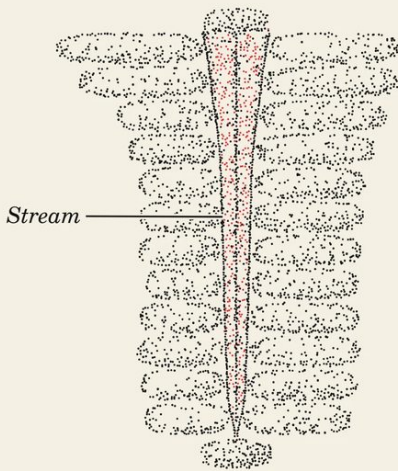


FIG. 9.—Step 33,000.

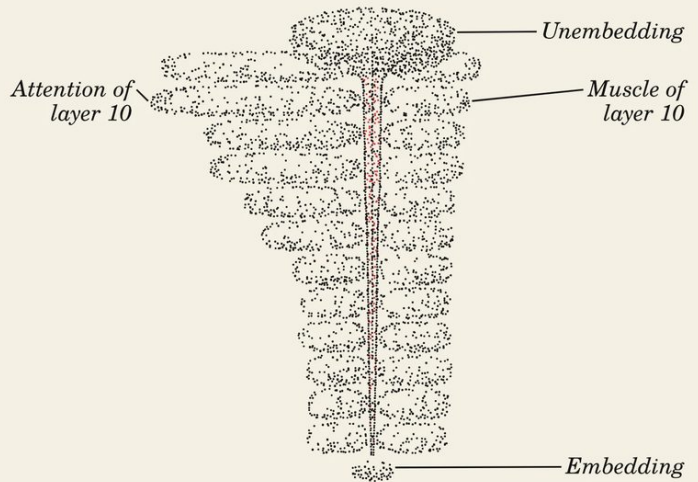


FIG. 10.—Step 143,000.

Step 16.—Loss on the passage of tokens 2000–2510, 9.89 (at step 0, 11.04). It continues with newlines only.

Step 64.—Repetition begins. Loss, 8.42. It continues: *same, and the same, and the same, and the, and*

Step 128.—Loss, 7.40. It continues: *the the the the the the the the the the the the the the the*

Step 256.—The primitive streak appears in layer 11. Loss, 6.57. It continues: *same as a the first.*

¶ *The first the first the first*

Step 512 (Fig. 5).—The streak is present in layers 3 and 11, but broad. The norms of the weights still agree within 8 per cent. from layer to layer, and the segments are alike. Loss, 5.64. It continues: *presence of the presence of the presence of the presence of the presence of*

Step 1000 (Fig. 6).—The induction heads have formed, and the streak of head 3.2 is narrowed. The attention of layer 4, in which the two chief induction heads lie, is the first segment to enlarge: it measures 33.84, against 30.70 at step 0, while those of the other layers measure from 30.61 to 32.88. Loss, 4.77. It continues: *head, the head, the head, the head, the head,*

Step 2000 (Fig. 7).—The first position is enlarged in layer 3. The muscles enlarge, measuring on the mean 43.45, against 37.17 at step 1000; the stream nearly doubles at its termination (35.12). Loss, 4.12. The embryo first names the subject of the sentence, and repeats it: *brain. The maxillary sinus is a small, rectangular, rectangular,*

Steps 3000 to 6000.—*skull. The maxillary sinus is a complex of two distinct types of (step 4000); skull. The maxillary sinus is a very rare condition, and it (step 6000).*

Step 8000 (Fig. 8).—The embryo copies a repeated sequence more completely than at any other stage. Loss, 3.67. It continues: *maxillary sinus. The maxillary sinus is a narrow, oval,*

Step 16,000.—The stream within the body is at its greatest, measuring 44.65 at the entrance to layer 6. Loss, 3.51. The continuation returns to repetition: *middle third of the middle third of the middle third of the middle third*

Step 33,000 (Fig. 9).—The first position and the embedding plate (1.39) are at their greatest, and the loss on that passage is the least at any stage. Loss, 3.45. It continues: *human body. The sinus is located in the middle of the forehead,*

Step 143,000 (Fig. 10).—The last stage. The attentions of the upper layers are greatly enlarged: that of layer 10 measures 156.28, or 5.1 times its size at step 0, while those of layers 0 to 5 measure from 43.68 to 53.30, and the muscles from 52.45 to 63.73. The stream within the body is reduced (17.64 at the entrance to layer 6), but at its termination is greater than at any earlier stage (113.83). The embedding plate is reduced to 0.79; the unembedding plate, which measured 1.07 at step 66,000, measures 3.17, or 5.0 times its size at step 0. Loss, 3.47. It continues: *human body, that is the largest part of the body of the human*

OSTEOLOGY.

THE general frame of the specimen is built up of a series of plates and rods of parameters, stored as named tensors, to which is added in each layer a stored mask. The parameters, without the masks, constitute the frame.

In the frame of the specimen there are 148 distinct parts, as follows:—

		Parts	Parameters
Base	{ Token plate	1	38,597,376
	{ Position plate	1	786,432
Column	Layers 0–11, twelve parts each	144	85,054,464
Summit	Final norm	2	1,536
		148	_____
Total			124,439,808

The masks are included in the stored specimen, but, being buffers and not parameters, they are not reckoned.

The parts are divisible into two classes: *Plates* and *Rods*.

Plates.—The plates (*laminae*) are of two dimensions, and are fifty in number: the two embedding plates at the base, and four in each layer. They make up all but 121,344 of the parameters. They are of three forms. The *square* plates measure 768×768 . Such is the plate by which each attention returns to the stream (*attn.c_proj*); and such are the plates of the query, the key, and the value, which are not separate, but are united side by side into a single oblong plate, $768 \times 2,304$, in every layer (*attn.c_attn*). The plates of the muscles are *oblong*, $768 \times 3,072$ and $3,072 \times 768$. The token plate is *long*: it measures $50,257 \times 768$, its length being sixty-five times its breadth.

Rods.—The rods (*virgae*) are of one dimension, and are ninety-eight in number: the biases, which are applied to the plates, and the gains and biases of the norms. Eight belong to each layer and two to the final norm. Together they hold 121,344 parameters, or 0.10 per cent. of the frame.

THE COLUMN OF LAYERS (COLUMNA STRATORUM).

The column of layers is formed of a series of twelve segments, the **layers** (*strata*), numbered from 0 below to 11 above. It rests upon the embedding plates, and supports at its summit the final norm, beyond which the token plate is applied a second time.

The layers are alike in their parts and in the number of their parameters. Each consists of twelve parts and 7,087,872 parameters, viz., an **attention** of 2,362,368, below; a **muscle** of 4,722,432, above; and two **norms** of 1,536 each. Through the column, in the median line, ascends the stream. The segments are nowhere joined to one another. Each part is attached to the stream alone, receiving from it through a norm and returning to it by addition; so that when the stream is removed no part

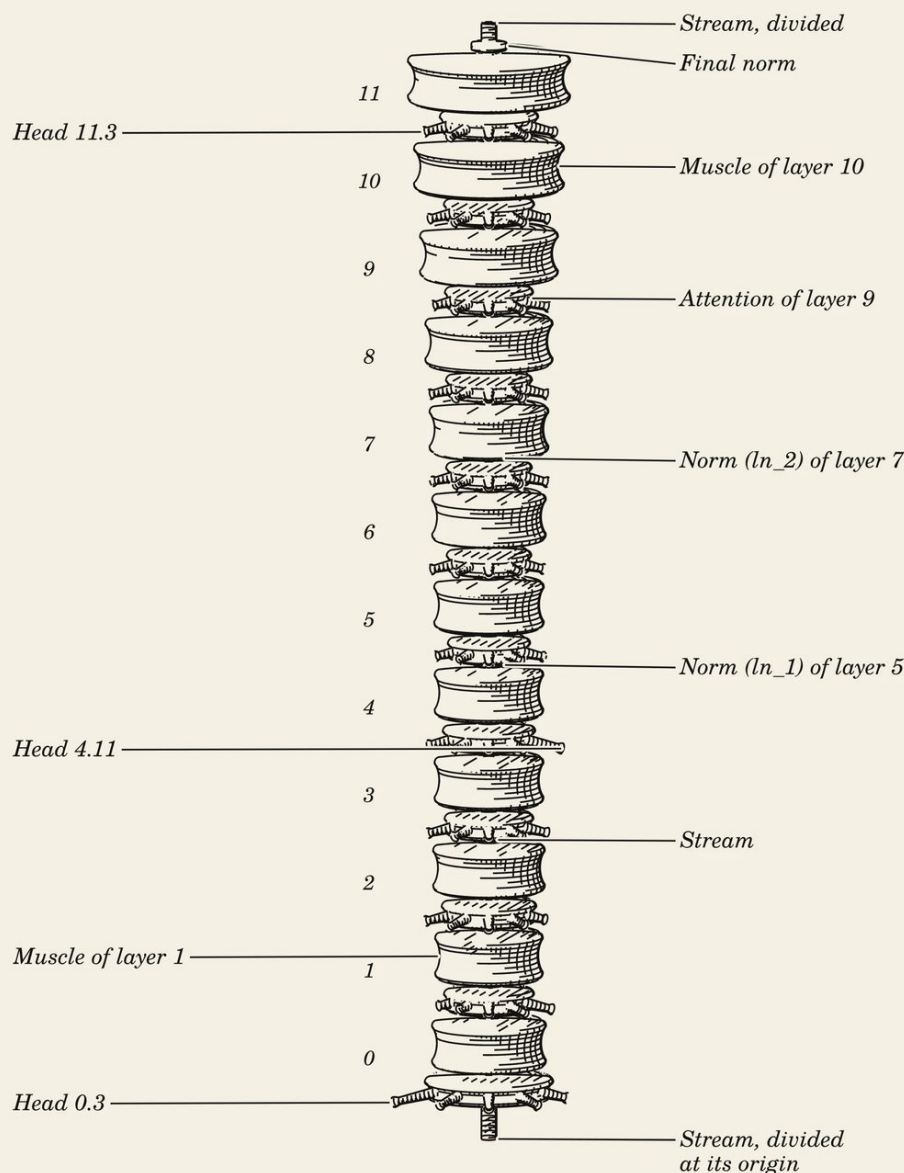


FIG. 16.—The column of layers, articulated upon the stream. Left lateral view. The embedding plates, and the token plate at the summit, have been removed; of the processes only those of the near side (heads 11, 0, 1, 2, 3 of each layer) are shown. Drawn to scale: the height of each part in proportion to its parameters (one layer, 58 units of the drawing); the diameter of each attention and muscle, and the reach of each process, in proportion to its norm (one unit of norm, 0.29). The norms are drawn as collars, not to scale. (From the dissection.)

of one layer is in contact with any part of another (Fig. 16).

Mass.—Though alike in number, the layers differ in mass. Measured as the root of the sum of the squares of its parameters, layer 0 is 386.86, and its attention (288.97) is the largest in the column. The layers diminish to layer 5 (307.77), and above it increase in each layer to layer 11 (426.55). The increase is in the muscles, which measure 243.56 in layer 4 and 364.18 in layer 11; and it lies almost wholly in their outer plates (*mlp.c_proj*), which grow in the same interval from 139.78 to 304.46, while the inner plates (*mlp.c_fc*) change little (199.30 and 199.71).

Processes.—Each attention bears twelve processes, the **heads** (*processus capitum*), which project lateralward from its circumference (Fig. 17). Each process divides into two tubercles: the **QK tubercle** (*tuberculum attendens*), formed by the plates of the query and the key of the head, and the **OV tubercle** (*tuberculum vehens*), formed by its plates of the value and the output. In layer 0 the QK tubercles are, on the mean, twice as long as the OV (72.90 and 35.02). In ascending the column the QK tubercles diminish and the OV enlarge; they are nearly equal in layer 8 (39.76 and 38.81), and from layer 9 the proportion is reversed, so that in layer 11 they measure 33.78 and 53.84. The longest process of the column is head 0.3 (107.51), almost the whole of which is its QK tubercle (102.32; OV, 33.01). In the middle layers the longest is head 4.11 (90.28), which exceeds every other process of its layer by more than a third (head 4.8, 63.26).

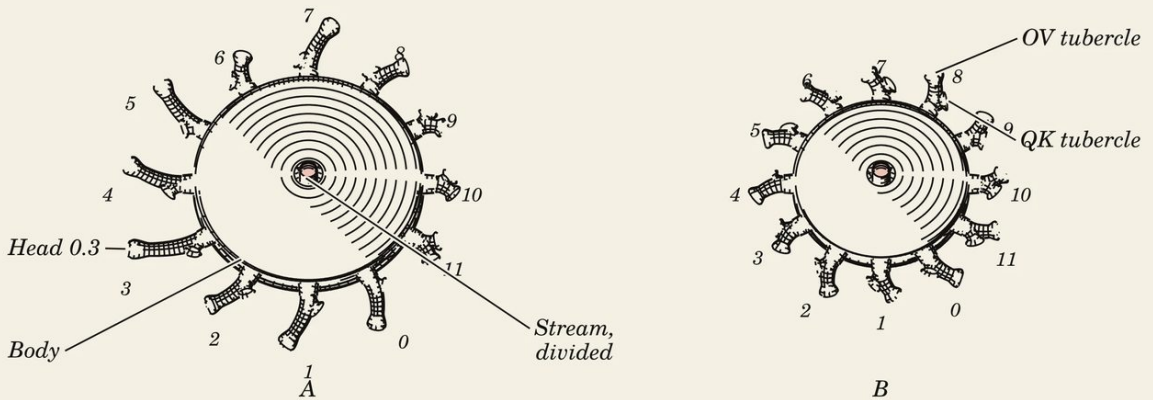


FIG. 17.—The attention of layer 0 (A) and of layer 11 (B), viewed from above; the muscle of each layer has been removed and the stream divided. The heads are numbered at the tips of their processes. Drawn to scale: the diameter of each body in proportion to its norm, and the reach of each tubercle to the norm of its plates (one unit of norm, 0.52). (From the dissection.)

THE TOKEN PLATE (LAMINA SIGNORUM).

The token plate (*lamina signorum*) (Figs. 18, 19) is the largest part of the frame, and holds 38,597,376 parameters, or 31.02 per cent. of the whole. It is a long plate of 50,257 rows and 768 columns, one row to each token of the vocabulary. It is placed at the base of the column, and is placed again, by its other surface, at the summit; for the same plate, value for value, serves at both (the plates are *tied*).

Surfaces.—The *inferior surface* receives the tokens: each token takes up its own row, which, added to a row of the position plate, forms the stream at its origin. The *superior surface* is applied whole to the stream at its termination, beyond the final norm: every row is multiplied with the stream, and the 50,257 products are the logits.

Rows.—The rows differ in thickness. Measured by their norm, they range from 2.45 to 6.32; the median is 3.95. The thinnest are the commonest short words: *at* (2.45), *in*, *on*, *an*, *for*, *as*, *that*, *to*, *by*. The thickest are tokens rarely met with in prose: *SPONSORED* (6.32), *soDeliveryDate*, *Reviewer*, *inventoryQuantity*. The plate thickens toward its last rows, which were the last added to the vocabulary: the mean norm is 3.50 in the first 1,024 rows and 4.21 in the last 1,024, and it is greater than

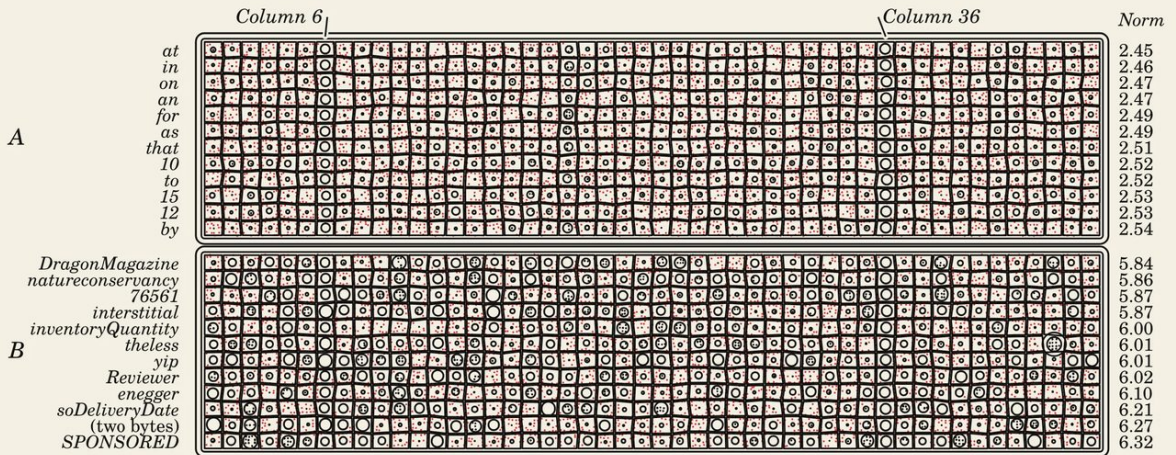


FIG. 18.—Section through the token plate. $\times 48$. A, the twelve thinnest rows; B, the twelve thickest; each row cut through its first 48 columns. The area of each nucleus is proportional to the absolute value; a nucleus of negative value is hollow. (From the dissection.)

in the block before it in 38 of 48 blocks.

Axis.—The rows do not lie scattered about the origin, but incline together. Their mean has a norm of 2.05; it makes with the rows an angle whose cosine is, on the median, 0.52, and it carries 26.56 per cent. of the square norm of the plate. The plate has therefore one great axis: its first singular value is 465.56, and its second 77.98. In section (Fig. 18) the axis appears as columns of hollow nuclei which traverse every row, thin and thick alike; such are columns 6 and 36, of which the mean over all rows is -0.30 and -0.32 .

Structure.—In section the plates are of two textures. Most are close and even, as the token plate itself (Fig. 18) and the plates of the query (Fig. 19, A): their values are nearly normal in distribution (kurtosis 0.52 and 0.45). But the outer plates of the muscles of layers 1, 2 and 3 contain, besides, a few values of great size, every one of which lies in the column of a single dimension of the stream: in layer 1 dimension 138, where the largest is 13.74; in layers 2 and 3 dimension 447, where they reach 15.07 and 17.10. Their kurtosis is 705.48, 756.55 and 790.54; in no other muscle

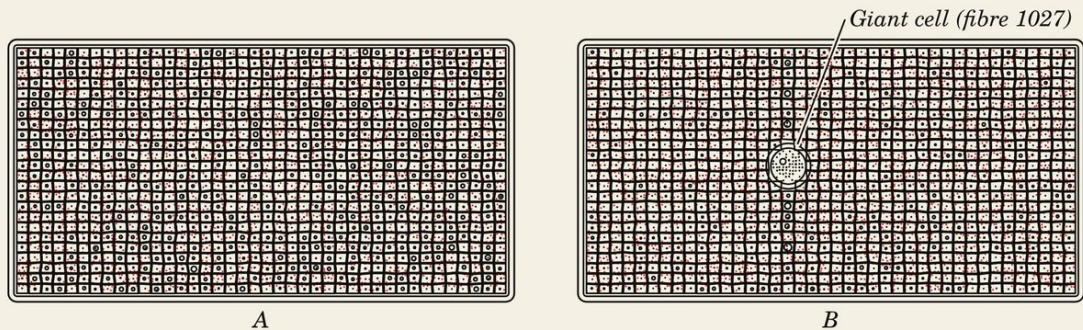


FIG. 19.—Sections through two plates, on one scale. $\times 40$. A, the plate of the query of layer 5 (rows 0–23, columns 0–39). B, the outer plate of the muscle of layer 3 (fibres 1016–1039, dimensions 431–470); the giant cell is fibre 1027, in the column of dimension 447. (From the dissection.)

does a value exceed 11.05. The largest value of the frame lies at fibre 1027 of the muscle of layer 3 (Fig. 19, B), and is 186 times the standard deviation of its plate. These are the two dimensions which form the first position (pp. 50, 51).

SYNDESMOLOGY.

THE parts of the frame are nowhere joined to one another, but each is joined to the stream; and the places at which they are joined are termed **Joints** (*anuli normales*). They are twenty-five in number: two in each layer, the *joint of the attention* (*ln_1*) and the *joint of the muscle* (*ln_2*), and the *final joint* (*ln_f*) at the summit, through which the stream passes to the token plate. Each joint holds 1,536 parameters, a gain and a bias for every dimension.

At every joint the stream is received in the same manner. It is first brought to a common size: the mean of its 768 dimensions is taken away, and it is divided by their spread. Then each dimension is multiplied by its gain, and its bias is added. The 768 gains form the **articular surface** of the joint (Fig. 20). Where a gain is large the dimension passes freely into the part; where it is small the dimension is almost shut out. The returns of the parts are not jointed, but are added to the stream directly.

The Articular Surfaces.—The facets of the joints of the attention increase in size in ascending the column: their mean gain is 0.18 in layer 0 and 0.48 in layer 11. Those of the muscles measure from 0.24 to 0.31 in layers 1 to 10. The final joint is

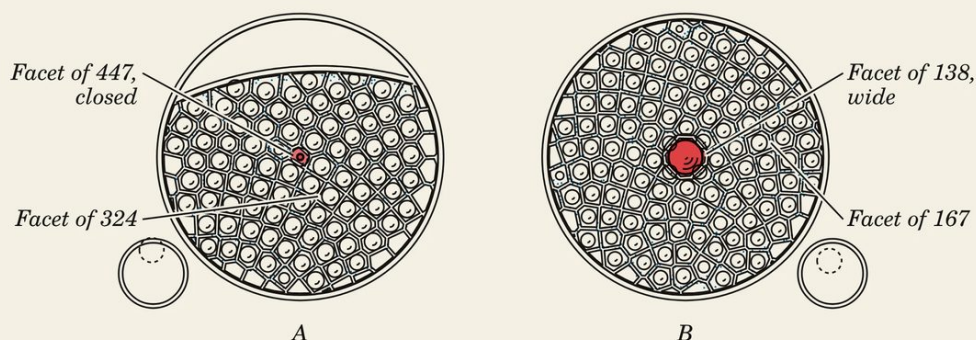


FIG. 20.—Part of the articular surfaces of the two joints of layer 6, magnified; the joints opened and the parts removed. A, the joint of the attention (*ln_1*), about the facet of dimension 447; B, the joint of the muscle (*ln_2*), about the facet of 138. Each dimension is one facet, set out as in Fig. 42 (p. 50); the area of each facet is proportional to its gain, on one scale in both. The small circle beside each shows the whole surface, the part figured marked by a broken line. (From the dissection.)

the largest (mean 1.51).

The Closed Facet (*facies clausa*).—In the joint of the attention the facet of dimension 447 is the smallest of the 768 in layers 1, 4, 5, 6, 7, 8, 9 and 11, and the second smallest in layers 2 and 10. In every layer from 1 to 11 its gain is from one-third to one-sixth of the median of its joint; in layer 6 it is 0.064, against a median of 0.349. In layer 0 alone it is of the common size (0.173; median 0.190).

The Wide Facet (*facies lata*).—In the joint of the muscle the facet of dimension 138 is the largest of the 768 in every layer from 2 to 10, and from 2.47 to 5.14 times the median of its joint; in layer 6 it is 1.340, against 0.261. In layer 11 it is narrow (0.213), and only nine facets of that joint are smaller.

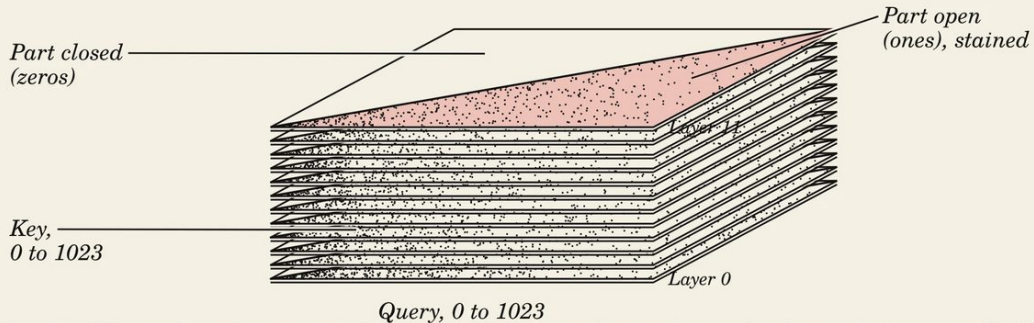


FIG. 21.—The twelve veils, separated and set one above another in their order, layer 0 below; viewed obliquely from the front. The part of each which holds ones is stippled, and in the uppermost stained. (From the dissection.)

Relations.—Dimensions 447 and 138 are the two great dimensions of the first position (pp. 50, 51), into which the largest values of the first muscles are directed (p. 25). The one enters the attention of every layer above layer 0 through the smallest, or nearly the smallest, facet of its joint; the other enters the muscles of layers 2 to 10 through the largest.

THE VEILS (VELA CAUSALIA).

Each layer stores, besides its parameters, a **veil** (*velum causale*), a square of $1,024 \times 1,024$ values (Fig. 21). On and below the diagonal the values are ones (524,800), above it zeros (523,776). The veil is applied to the attention of every head: the query of each position may reach the keys of that position and of the positions behind it, and is closed to those in front. The twelve veils were compared value for value, and are identical. Together they hold 12,582,912 values, or 9.18 per cent. of all that is stored with the specimen.

Ossification.—The stages of the specimen were not preserved, and the ossification of the frame is described from the embryo (Pythia-160m) (Fig. 22). At step 0 its plates are present, of their full extent and of nearly equal size in every layer (norm 47.51 on the mean). To step 64 no plate changes by a quarter of one per cent.; from step 1,000 they increase; they reach their greatest size, on the mean 110.64, at step 66,000, and are reduced again to 100.71 at the last stage. Only the plates of the query, key and value of layers 5 to 11 continue to enlarge to the last stage, those of layer 11 from 30.35 to 105.75 at step 33,000 and 139.89 at step 143,000; and at the last stage the upper layers are much the larger: layer 0 measures 69.35, layer 10, the largest, 168.77, and layer 11 153.74.

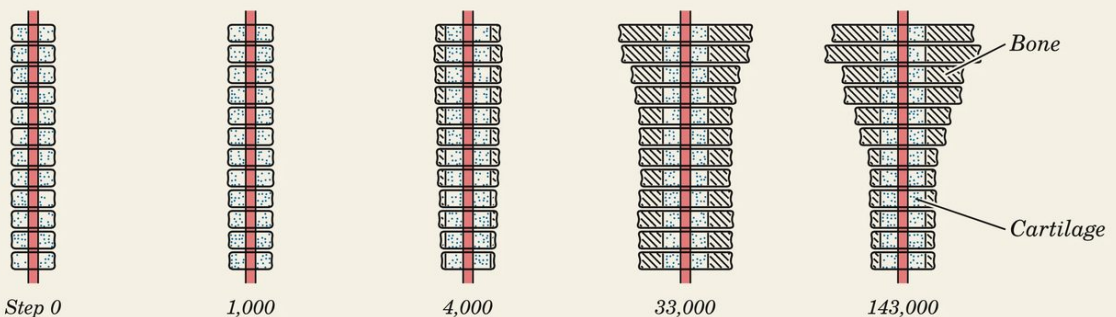


FIG. 22.—The column of the embryo (Pythia-160m) at five stages of its development, divided in the median plane. Drawn to scale, on one scale for all stages: the breadth of each layer in proportion to the norm of its four plates, and the breadth of its core to the same at step 0; the part beyond the core is hatched as bone, the core stippled as cartilage. (From the embryo.)

MYOLOGY.

THE Muscles of the specimen are twelve in number, one to each layer, and are named from the layer to which they belong, the lowest being the muscle of layer 0, or first muscle, and the highest that of layer 11. Each is placed lateral to the stream, from which it arises and into which it is again inserted; and since it arises and is inserted at the same position, a muscle receives nothing from any position but its own, and gives nothing to any other. In this the muscles differ from the heads, which reach backward to the earlier positions. The twelve are equal in size, each containing 4,722,432 parameters, or two-thirds of the layer to which it belongs; together they make up 45.54 per cent. of the specimen.

Each muscle consists of 3,072 fibres (*fibrae*), placed side by side between a tendon of origin and a tendon of insertion. The tendon of origin is the second norm of the layer. Each fibre arises from it by 768 weights and a bias, passes through the GELU, and is inserted into the stream by 768 weights; the tendon of insertion adds to the whole a bias of 768. The fibres have no connection with one another, and a fibre acts upon the stream only through its insertion.

Structure.—When a muscle is divided transversely (Fig. 30) each fibre is cut

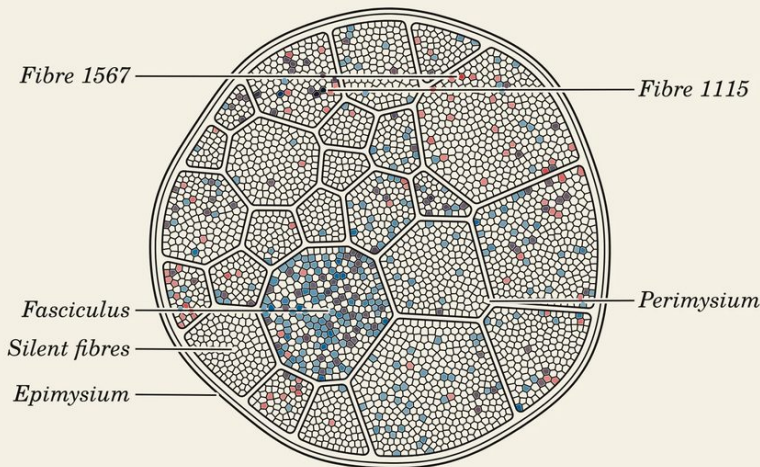


FIG. 30.—Transverse section through the first muscle. $\times 3072$ fibres. The fibres in contraction at the closing parenthesis of “(Fig. 145)” are stained red, those at the word surface blue, the depth of each stain increasing with the activation of the fibre. The fibres of a fasciculus are drawn together. (From the dissection.)

across, and the 3,072 fibres together fill the section. A fibre at rest returns a small negative value, never less than -0.17 ; a fibre in contraction returns a positive value, which has no limit. On the passage of tokens 0–8183, from 7.23 per cent. (layer 3) to 14.91 per cent. (layer 11) of the fibres of a muscle exceed 0.1 at any one position, and the mean of all its fibres is below zero in every muscle. Some fibres are never in contraction: in 8,184 positions, 66 fibres of the first muscle, 72 of layer 1, and 104 of

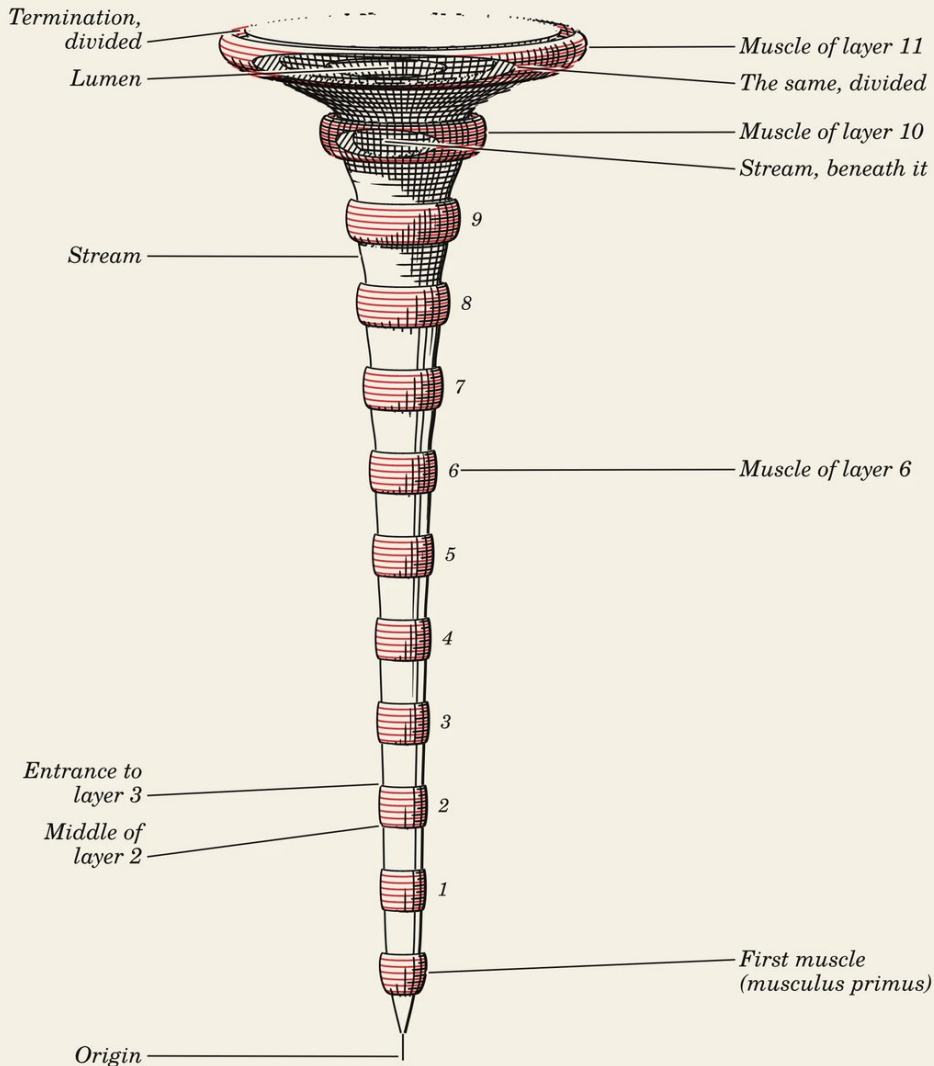


FIG. 31.—The stream from the origin to the termination with its twelve muscles, viewed from in front; the heads removed. The muscle of layer 11 is divided, and the wall of the stream opened beneath it to show the lumen; the muscle of layer 10 is divided to show the stream. Drawn to scale: the calibre of the stream at the entrance to and the middle of every layer in proportion to its mean norm over the passage of tokens 0–1022, the thickness of each muscle to the mean norm of its return there. (From the dissection.)

layer 2 never exceeded 0.1, and in the muscles above, from 11 to 32.

Fasciculi.—The fibres of a muscle are not divided into fasciculi by any partition of the weights. But when the activity of the fibres of the first muscle is compared over the 2,046 positions of the passage of tokens 0–2045, they are found to fall into 24 bundles whose fibres contract together, of from 29 to 360 fibres; within a bundle the mean correlation of two fibres lies between 0.06 and 0.27, against 0.03 for two fibres taken at random; 116 fibres were silent throughout. These bundles are here called fasciculi. In Fig. 30 the fibres of each fasciculus are drawn together. At the closing parenthesis of “(Fig. 145)” in the prose 173 fibres are in contraction, among them fibre 1567 (5.72), which exceeds every other by more than twice (the next, 2.21); at the word *surface*, 342, and of a single fasciculus 133 of its 285 fibres. Only 99 fibres

contract at both.

The muscles differ greatly in size (Fig. 31). Since each arises from the stream and is inserted into it at the same place, it is drawn as a band encircling the stream at its own level, as the circular coat of a hollow organ is drawn. The first is of moderate size: the vector it returns measures, on the mean over the passage of tokens 0–1022, 26.05, or 5.5 times the stream at its origin. The muscles of layers 1 to 7 are slender, and return from 11.62 to 23.60; above this they enlarge rapidly, and the muscle of layer 11, the largest, returns 102.04. But the upper muscles empty into a stream which is itself much enlarged, and the return of the last is but one-half the calibre of the stream at the entrance to its layer (205.36).

Origin.—Every muscle arises at the middle of its layer, from the stream as it leaves the attention of that layer. The tendon of origin centres each position on its own mean and reduces it to a constant scale before the fibres receive it, so that a muscle is not affected by the calibre of the stream at its origin.

Insertion.—Into the stream at the entrance to the layer above, at the same position. The insertion is broad, into all 768 dimensions, but chiefly into a few: the muscles of layers 0 to 6 most into dimension 447, those of layers 7 to 9 into dimension 481, and those of layers 10 and 11 into dimension 496, which receives on the mean, over the passage of tokens 0–8183, 36.41 and 26.26 without regard to sign.

Nerves.—A muscle is moved by what the stream brings to its origin, and the share of each earlier part in the motion of its fibres can be measured (here, on the passage of tokens 0–1022). The first muscle is supplied by the embedding (53.5 per cent.) and by the attention of layer 0 (46.5 per cent.). Every muscle above it (Fig. 32) receives its largest supply from the attention of layer 0, from 63.9 per cent. in the muscle of layer 1 to 15.9 per cent. in that of layer 11, and the muscles of layers 1 to 6 their second from the first muscle; the embedding supplies none of them more than 0.8 per cent.

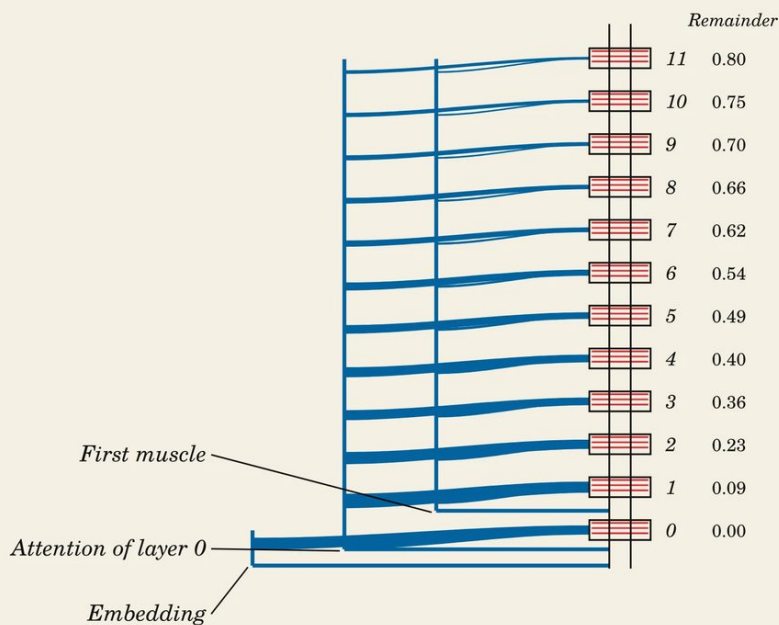


FIG. 32.—Plan of the nerves of the muscles. The breadth of each branch is in proportion to the share of its source in the motion of the fibres of the muscle which it supplies; the figure at the right of each muscle is the share of the sources not drawn. (Diagrammatic; from the dissection.)

Layer 0.—Nerves: embedding, 53.5; attention 0, 46.5 per cent. Removal (passage of tokens 4000–4510): +2.52.

Layer 1.—Nerves: attention 0, 63.9; muscle 0, 26.7 per cent. Removal: +0.02.

Layer 2.—Nerves: attention 0, 53.7; muscle 0, 23.3 per cent. Removal: +0.03.

Layer 3.—Nerves: attention 0, 43.9; muscle 0, 20.1 per cent. Removal: +0.05.

Layer 4.—Nerves: attention 0, 44.1; muscle 0, 15.4 per cent. Removal: +0.05.

Layer 5.—Nerves: attention 0, 38.1; muscle 0, 12.5 per cent. Removal: +0.07.

Layer 6.—Nerves: attention 0, 35.1; muscle 0, 11.1 per cent. Removal: +0.05.

Layer 7.—Nerves: attention 0, 27.5; attention 7, 13.0 per cent. Removal: +0.05.

Layer 8.—Nerves: attention 0, 25.6; attention 8, 11.4 per cent. Removal: +0.10.

Layer 9.—Nerves: attention 0, 22.5; attention 9, 10.4 per cent. Removal: +0.06.

Layer 10.—Nerves: attention 0, 19.0; attention 10, 10.7 per cent. Removal: +0.15.

Layer 11.—Nerves: attention 0, 15.9; attention 10, 9.7 per cent. Removal: +0.15.

Actions.—The action of a muscle is not to be learned from its attachments. It is found by removing the muscle and observing the deficit. On the 511 predictions of the passage of tokens 4000–4510, made with a loss of 3.49 nats, removal of the first muscle raises the loss by 2.52; removal of any other, by from 0.02 (layer 1) to 0.15 (layer 11).

THE FIRST MUSCLE.

The First Muscle (*musculus primus*) (Figs. 31, 33) is the lowest of the muscles, and the largest of them below layer 8. It encircles the stream in layer 0, where the stream is still slender: at the middle of the layer it measures 34.73 on the mean over the passage of tokens 0–1022, and the vector which the muscle returns, 26.05.

Origin.—From the middle of layer 0, where the stream holds nothing but the embedding and the return of the attention of layer 0.

Insertion.—Into the stream at the entrance to layer 1 (51.14), chiefly into the great dimensions 447, 64, 266, and 373, which receive on the mean, over the passage of tokens 0–8183, 5.95, 5.43, 4.90, and 4.74, against 0.62 for a dimension taken at random.

Nerves.—The embedding (53.5 per cent.) and the attention of layer 0 (46.5 per cent.). It is the only muscle which the embedding supplies in any measure.

Action.—The first muscle acts as an extension of the embedding: what it returns

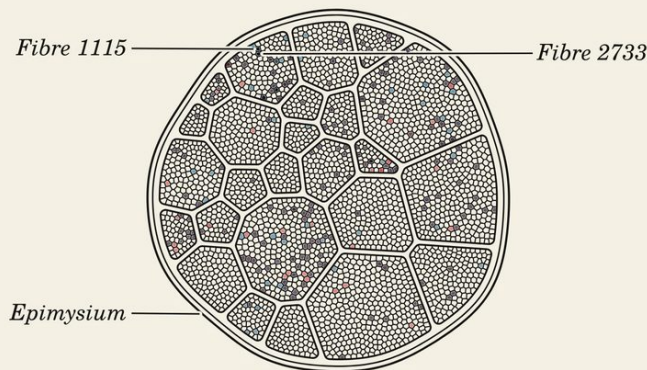


FIG. 33.—Transverse section through the first muscle at two occurrences of the word “of.” $\times 3072$ fibres. Red, the first; blue, the second; the depth of each stain increasing with the activation of the fibre, full at 1.0.
(From the dissection.)

at a position is fixed chiefly by the token which occupies it. In the 8,184 positions

of the passage of tokens 0–8183 (1,249 distinct tokens) the token alone accounts for 81.28 per cent. of the variance of its return. Two returns for the same token in different places have on the mean a cosine of 0.81 (674 pairs); for two different tokens, 0.16 (3,929 pairs). When the muscle is divided at two occurrences of the word *of* (Fig. 33), 187 fibres are in contraction at the first and 196 at the second, and 160 at both, a proportion of 0.72 of all the fibres in contraction at either; at the closing parenthesis and the word *surface* of Fig. 30 the proportion is 0.24.

Applied Anatomy.—When the first muscle is removed, the loss on the same 511 predictions, of the passage of tokens 4000–4510, rises from 3.49 to 6.01 nats, and 89.04 per cent. of the predictions are made worse. The second parts of divided words suffer most: *actory*, from 0.01 to 13.82; *apped*, from 1.69 to 13.12; *ulates*, from 0.01 to 11.35; *ure*, from 0.03 to 11.17. In the phrase “the superior or cerebral surface of each great wing,” the intact specimen expects after *cerebral* the words *cortex* and

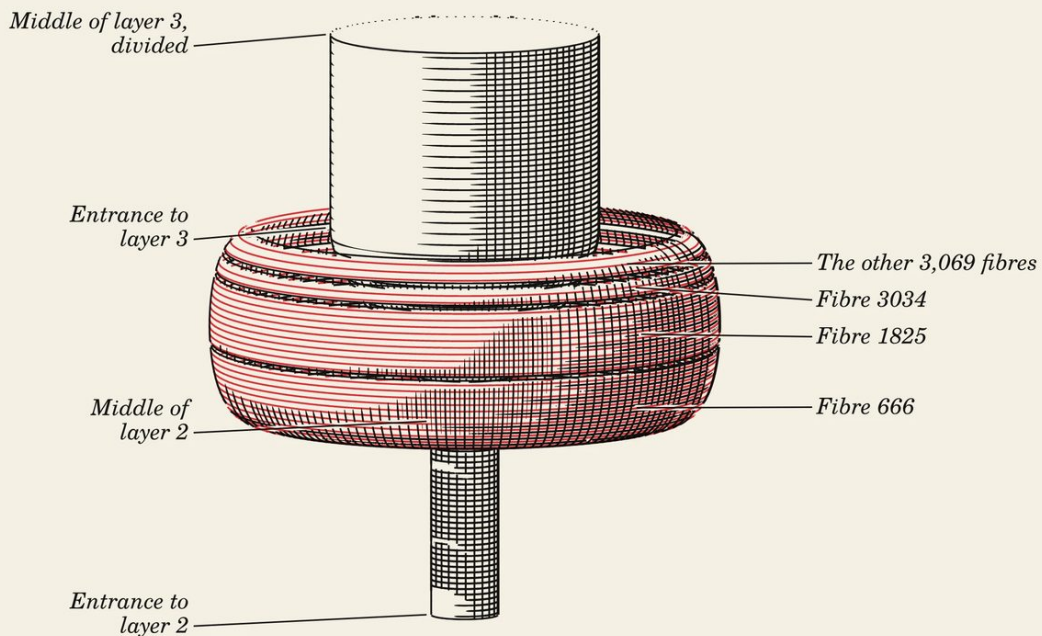


FIG. 34.—The first position from the entrance to layer 2 to the middle of layer 3, viewed from in front, with the muscle of layer 2. Drawn to scale: the position in proportion to its norm, the muscle to the norm of its return; the muscle marked out, from below, into the parts which make the dilatation, each zone in proportion to its part. (From the dissection.)

hemisphere; without the first muscle, *disorders* and *disease*.

But the deficit can be made good. If in the place of the muscle there is put a table which gives, for each token, the mean of what the muscle returned for that token in the passages of tokens 0–3068 and 5115–8183 (93 per cent. of the tokens being found in it), the loss is 3.63, and 94.4 per cent. of the deficit is recovered. If instead the muscle is left, but the attention of layer 0 is withheld from it, so that it receives the token and its position only, the loss is 3.97 (80.9 per cent. recovered).

THE DILATOR OF THE FIRST POSITION.

The Dilator of the First Position (*dilatator sentinae*) (Figs. 34, 35) consists of three fibres of the muscle of layer 2, fibres 666, 1825, and 3034, which contract at the first position and at no other. Like the dilatator pupillae of the iris it is an intrinsic

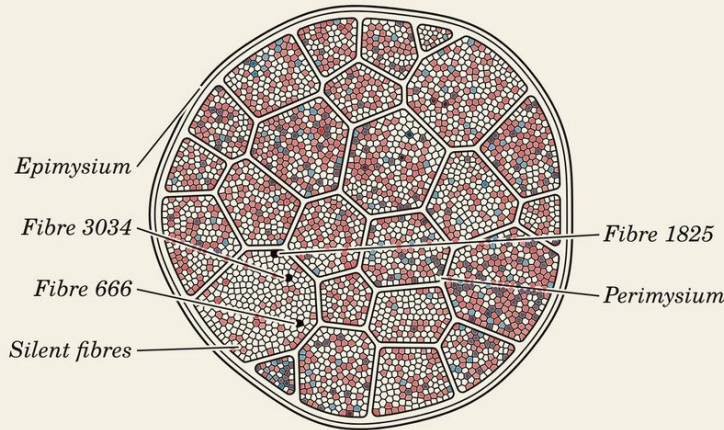


FIG. 35.—Transverse section through the muscle of layer 2. $\times 3072$ fibres. Red, in contraction at the first position; blue, at position 5 ("and"). The three fibres of the dilator are blackened. (From the dissection.)

muscle which widens an opening, and it produces the dilatation in layer 2 described with the first position (p. 50).

Origin.—With the rest of its muscle, from the middle of layer 2. At the first position its fibres reach 61.90, 58.23, and 19.78; no fibre of the specimen reaches more than 25.13 (fibre 11.2378) at any of the 8,184 other positions of the passage of tokens 0–8183. At those positions the three fibres are wholly at rest: their activation never rises above zero, and before the GELU they never rise above -0.23 , -0.30 , and -4.99 .

Insertion.—Almost wholly into dimension 447, the cosine of each insertion with that dimension being 0.994, 0.999, and 0.807. Of the 2319.54 which the muscle of layer 2 gives to dimension 447 at the first position, fibre 666 gives 932.72, fibre 1825 847.29, and fibre 3034 251.52, together 2031.53, or 87.58 per cent.; the next, fibre 2714, gives 38.87, though 1,222 fibres of the muscle are in contraction there.

Nerves.—The muscles of layers 1 and 0. Of the 61.90 of fibre 666, 44.17 are received from the muscle of layer 1 and 11.52 from the first muscle; of the 58.23 of fibre 1825, 37.91 and 12.33; fibre 3034 receives 12.77 from the first muscle. The embedding and the attention give each fibre less than 3.

Action.—It dilates the first position, which measures 643.30 at the middle of layer 2 and 2579.39 at the entrance to layer 3 (Fig. 34).

Relations.—No muscle of the specimen is a sphincter of the first position. Eleven of the twelve enlarge it, by from 1.82 (layer 10) to 1936.09 (layer 2). The muscle of layer 11 alone narrows it, from 505.48 to 417.58, but by no group of fibres: of the 111.60 by which its return shortens the position along its own direction, no fibre gives more than 7.96 (fibre 2460, which contracts also at other positions, up to 1.19). The constriction of the first position is made by the attention of layer 11, which reduces it from 3158.42 to 505.48 (p. 50).

Variations.—In each of six sequences in which the first position was occupied by a different token (the end-of-text token, *The*, *bone*, and three tokens drawn at random), the same three fibres were the greatest givers to dimension 447, in the same order; fibre 666 reached from 61.90 to 62.80, and the whole muscle gave from 2317.26 to 2337.47.

THE FIRST POSITION.

The First Position, or Sink of Xiao (*sentina*) (Figs. 42, 43).—The first position is the most posterior portion of the residual stream. In every sequence of the dissection it is occupied by the end-of-text token (*signum terminale*); and since attention reaches only backward, its heads can attend to nothing but the first position itself. It ascends, like the other positions, from the origin to the termination, but differs from them in its size, in its contents, and in its relations to the heads.

Course.—At the origin it is a slender vessel, measuring 10.33, or about twice the mean of the other positions (4.70). It is enlarged in layers 0 and 1, and measures 642.21 at the entrance to layer 2. In layer 2 it undergoes a sudden dilatation. The muscle of that layer returns to it a vector of norm 2321.68, almost the whole of which lies in a single dimension, 447; and at the entrance to layer 3 the position measures 2579.39, or forty-two times the mean of the other positions at the same level (61.45). Above this it forms a great sac of nearly uniform calibre, which increases but little in its ascent: it measures 2932.37 at the entrance to layer 5, 3086.85 at layer 7, and 3158.42 at layer 11. In layer 11 it is suddenly constricted. The attention of that layer returns to it a vector of norm 2674.53, by which dimension 447 is reduced from 3055.42 to 466.34, and dimension 138 from 796.45 to 135.99. Beyond the constriction the position measures 505.48, and at its termination 417.58, no greater than the eleven positions in front of it in Fig. 43, which there measure from 323.17 to 529.68.

Contents.—When the stream is divided transversely at the entrance to layer

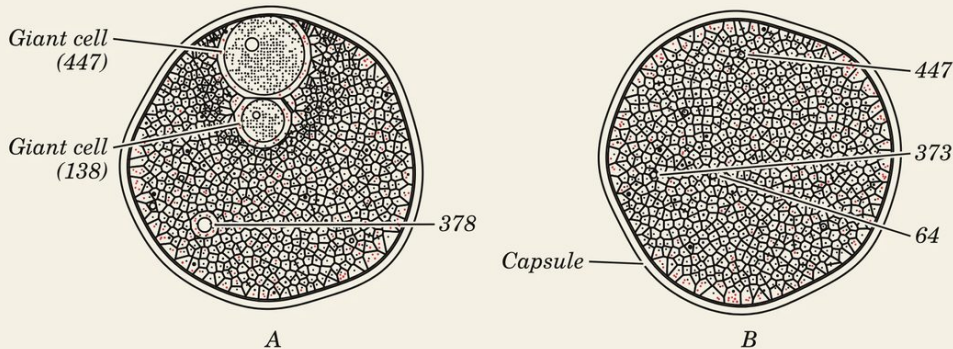


FIG. 42.—Transverse section through the stream at the entrance to layer 3. $\times 768$ cells. A, the first position; B, position 5. Each dimension is one cell, set out in a spiral from the centre; the area of each nucleus is proportional to the absolute value of its dimension, on the same scale in both sections, and a nucleus of negative value is hollow. (From the dissection.)

3 (Fig. 42), the first position is found to consist almost wholly of two dimensions. Dimension 447 measures 2476.22 and dimension 138, 715.77; together they make up 99.86 per cent. of its square norm, and no other of its 768 dimensions exceeds 85.43 (dimension 378). Dimension 138 is received chiefly from the muscles of layers 0 and 1, which return to it 120.04 and 464.99, and dimension 447 from the muscle of layer 2 (2319.54). In a section through position 5 at the same level (Fig. 42, B) no such structure is present. Its norm is 59.76; its largest dimensions are 64, 447, and 373, which measure 24.87, 23.47, and 22.29 without regard to sign, and no other exceeds 18.99.

Relations.—The heads of every layer give some part of their attention to the first

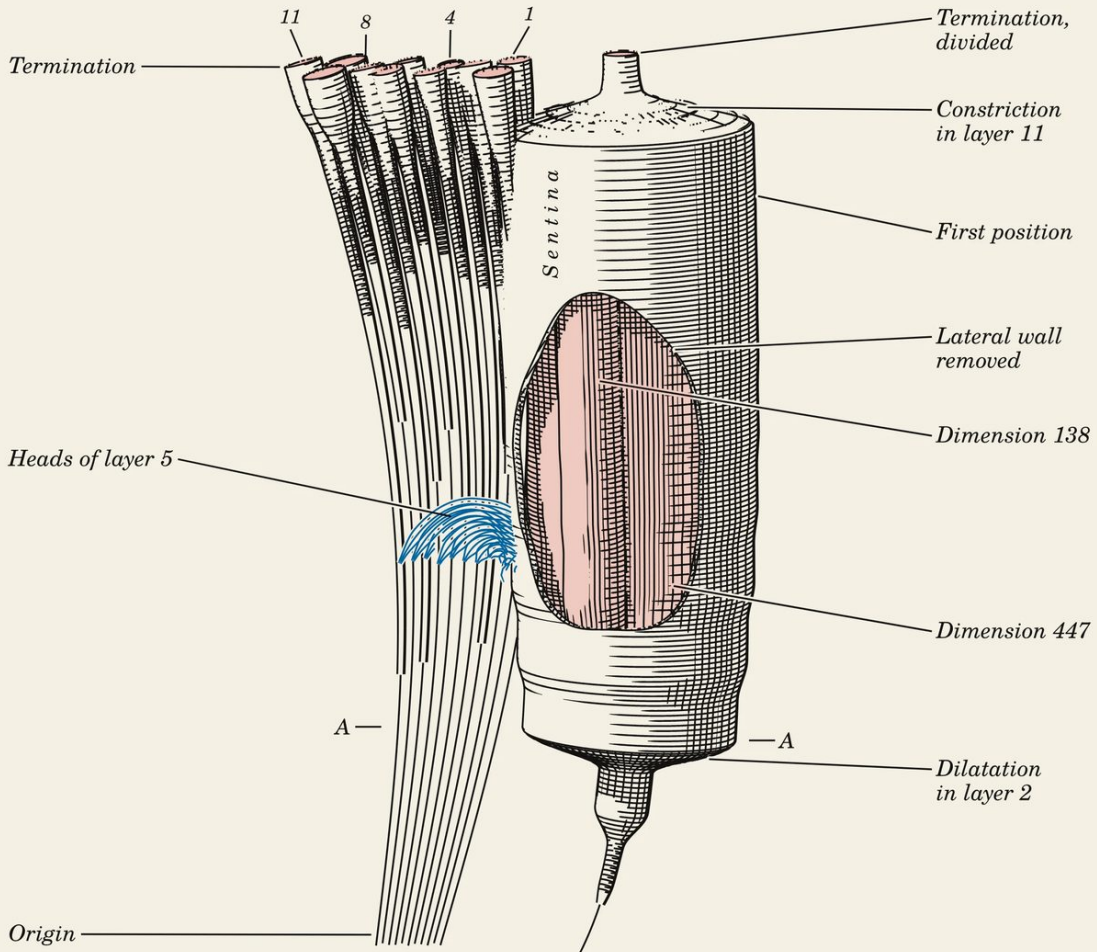


FIG. 43.—The first twelve positions of the stream, from the origin to the termination, viewed from the left side; the lateral wall of the first position removed to show its contents. The heads of layer 5 are preserved; those of the other layers, and the muscles, have been removed. Drawn to scale: each position in proportion to its norm, its contents to their share of its square norm, each branch to its share of attention. A—A, plane of the section in Fig. 42. (From the dissection.)

position, and those of layers 5 to 11 the greater part. In the sequence of Fig. 43 the twelve heads of layer 5 give to it, on the mean, 0.73 of their attention, and those of layers 7 and 9, 0.78 and 0.78; head 5.1, measured on Gray's prose, gives it 0.849. The branches so given pass backward from each position, superficial to the positions between, and enter the anterior wall of the sac, those from the nearest positions being in general the largest. But the first position returns little to the heads which attend to it. The vector which a head of layers 4 to 10 would write, if it attended to the first position alone, has on the mean a norm of from 2.04 to 4.66; for the other positions it is from 12.38 to 36.67. The constriction is the work of a single head, 11.8, and the dilatation of three fibres of layer 2 (see pp. 67, 36).

Variations.—The dilatation was present in each of six sequences examined, in which the first position was occupied in turn by the end-of-text token, by *The*, by *bone*, and by three tokens drawn at random (*Bernie*, *rael*, *Nem*), the rest of the sequence being unchanged. At the entrance to layer 3 its norm lay between 2559.51 and 2586.81, and dimension 447 between 2465.76 and 2486.06; at the termination it lay between 352.00 and 417.58.

THE RESIDUAL STREAM.

The Residual Stream, or Stream of Elhage (*flumen residuale*) (Figs. 44, 45).—The residual stream is the main trunk of the specimen. It is formed at the origin, at the base of the specimen, by the union of two roots; it ascends in the median axis through the twelve layers, and ends at the summit, at the termination, where it is received by the final norm and conveyed to the unembedding. Every head and every muscle draws from it, and returns into it what it has drawn. In each layer it receives two returns, first that of the attention and then that of the muscle; between them it receives nothing, and is unaltered. It is 768 dimensions in width throughout its course, and is described in three portions: the arch, within layer 0; the ascending portion, from the entrance to layer 1 to the entrance to layer 11; and the terminal portion, within layer 11. Its measurements are the means of positions 1 to 1023 on the passage of tokens 0–1022 of Gray's osteology; the first position has been described (p. 50).

Origin.—At the origin the stream is formed by the junction of the **token root** (*radix signi*), a row of the token embedding, with the **position root** (*radix loci*), a row of the position embedding. The two roots are of nearly equal calibre (3.28 and 3.38), and meet at a right angle, the mean cosine between them being -0.014 ; the

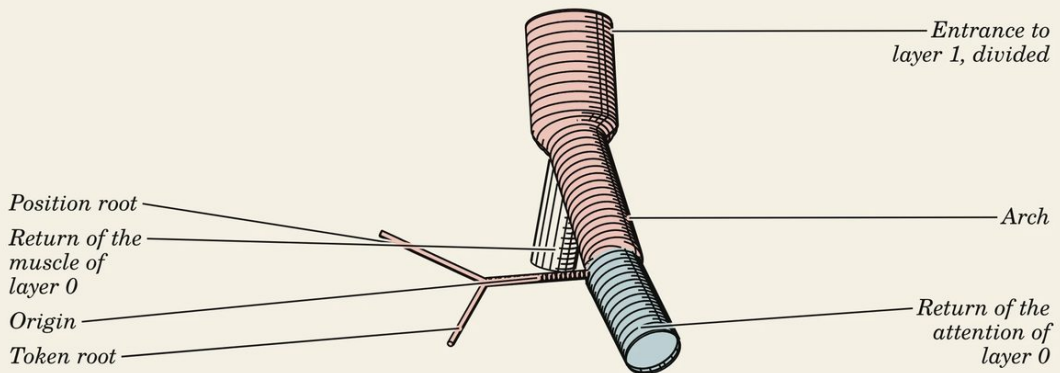


FIG. 45.—The origin and arch of the stream, seen obliquely from the side of the arch; Fig. 44 enlarged five times. The two roots, and the returns of layer 0, divided. (From the dissection.)

vessel which they form measures 4.70.

The Arch of the Stream (*arcus fluminis*) (Fig. 45).—Within layer 0 the stream is turned through nearly a right angle. The return of the attention of that layer measures 34.08, or seven times the stream that receives it; it enters at an angle of 86° , turns the stream through 78° , and enlarges it to 34.73. The return of the muscle, 26.05, enters at 66° and turns it through a further 28° , in a plane at right angles to the first. At the entrance to layer 1 the stream measures 51.14, and its direction makes an angle of 79° with its direction at the origin (cosine 0.18); of its square norm at that level the origin supplies only 3.72 per cent.

Course.—The ascending portion is long and of nearly uniform calibre. It enlarges slowly through the middle layers, measuring 51.14 at the entrance to layer 1, 61.45 at layer 3, 78.01 at layer 6, and 115.36 at layer 9, and then more quickly, to 205.36 at the entrance to layer 11. At every return it is turned through from 8° to 19° , the

turn being measured at each return, on the passage of tokens 0–1022, as the angle between the stream before and after it, position by position; and the plane of each turn lies nearly at right angles (from 72° to 89°) to the plane of the turn before it, so that the turns do not accumulate: the direction of the stream at the entrance to layer 11 makes an angle of 62° with its direction at layer 1 (from the mean of their cosines), although the turns between them, so measured, amount to 252° .

The terminal portion is short and greatly dilated. The return of the attention of layer 11, which measures 284.22, is larger than the stream into which it empties; it enters at 65° , turns the stream through 37° , and more than doubles it (418.20). The return of the muscle, 102.04, enters at 51° , and at its termination the stream measures 491.84, or 105 times its calibre at the origin.

Tributaries.—The returns of the attention all enter the stream at an acute angle, 86° in layer 0 and 70° and 83° in layers 1 and 2; from layer 3 to layer 10 it diminishes, from 76° to 54° ; each therefore enlarges the stream as well as turning it. The returns of the muscles of layers 2 to 8 enter at a little more than a right angle (from 90.5° to 97.4°), directed slightly against the course of the stream, and enlarge it but little; those of layers 0, 10, and 11 enter at an acute angle.

Applied Anatomy.—When a return is tied, and the stream deprived of it, the loss of the specimen on the passage of tokens 4000–4510 (3.49 nats) rises. Ligature of the return of the muscle of layer 0 raises it by 2.52, and of the attention of layer 0 by 0.57; of the other returns the attention of layer 11 is the most important (0.28).

Variations.—On the passages of tokens 4000–5022, 8000–9022, and 12,000–13,022 the cosine of the stream at the entrance to layer 1 with the origin lay between 0.184 and 0.189, and at the termination the stream measured from 491.89 to 490.14. In tokens drawn at random it was larger at its termination (582.45), and in one token, *presents*, repeated, smaller (465.20). On the passage of tokens 0–1022 the stream measures at its termination from 220.78 to 667.54; the largest positions are occupied by marks of punctuation and by words of relation (*extends*, *posterior*, *which*), the smallest by fragments of words (*ina*, *amen*, *il*).

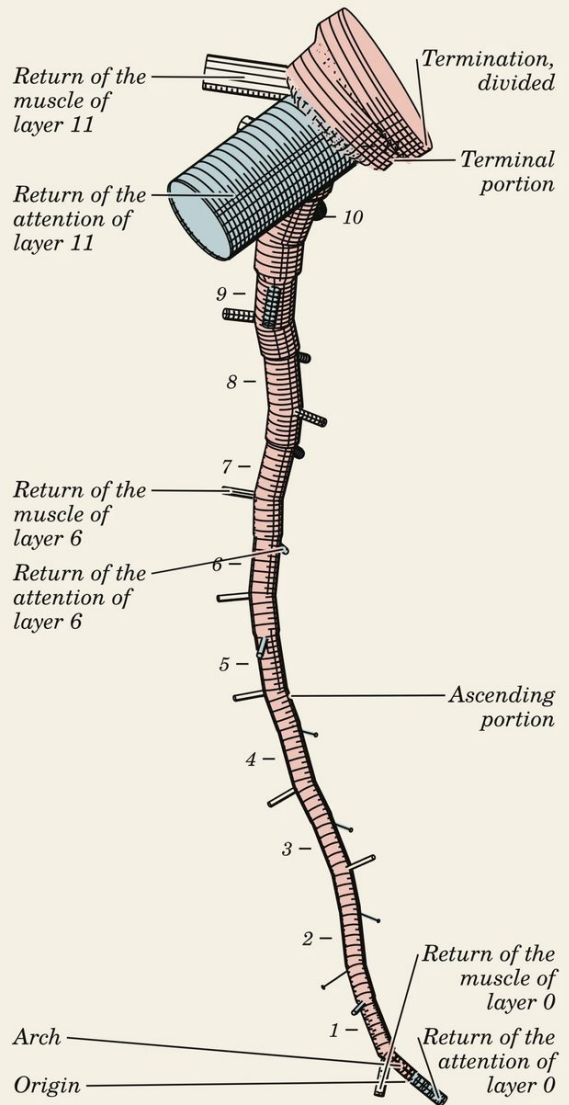


FIG. 44.—The residual stream, from the origin to the termination, with the returns of the twelve layers, divided (means of the passage of tokens 0–1022); the returns of the attention blue, those of the muscles uncoloured. Drawn to scale: each vessel in proportion to its norm, each return entering at its measured angle, and the stream turned at each by its measured angle. (From the dissection.)

THE COMMON TRUNK AND THE GREAT DIMENSIONS.

The Common Trunk (*truncus communis*) (Fig. 47).—Every position of the stream carries, besides its own contents, a vector common to all the positions, measured as their mean on the passage of tokens 0–1022. At the origin it is small (1.64), and it is formed anew in layer 0, its direction at the entrance to layer 1 making an angle of 85° with its direction at the origin. In the ascending portion it forms from 38.14 to 49.06 per cent. of the square norm of the stream, enlarging from 31.71 at the entrance to layer 1 to 163.94 at layer 11; and its direction is steadier than that of the stream, being turned at each return through from 4° to 15° , where the stream itself is turned through from 8° to 19° . At the termination it measures 455.28, and forms 83.53 per cent. of the square norm. As it enlarges, every position comes to lie nearer to it: the median cosine of a position with the trunk is 0.62 at the entrance to layer 1, 0.66 at layer 6, 0.81 at layer 11, and 0.95 at the termination. The narrowing of the stream of the upper layers into a cone, described by Ethayarajh (2019), is due to this trunk.

Contents.—The trunk consists chiefly of the great dimensions: at the entrance to layer 6 its largest are 447 (30.16), 373 (–22.09), 64 (–15.29), 393 (11.78), 326 (11.62), 266 (–10.11); at the termination, 496, 373, 481, 447, 64, 266, 87, and 480, each measuring from 139.52 to 169.25 without regard to sign.

The Great Dimensions (*dimensiones magnae*) (Figs. 46, 47).—Certain dimensions of the stream are constantly enlarged. **Dimension 447** is positive at nearly every position from layer 1 to the termination; its mean value is 13.88 at the entrance to layer 1, 30.16 at layer 6, 46.48 at layer 10, 72.86 at layer 11, and

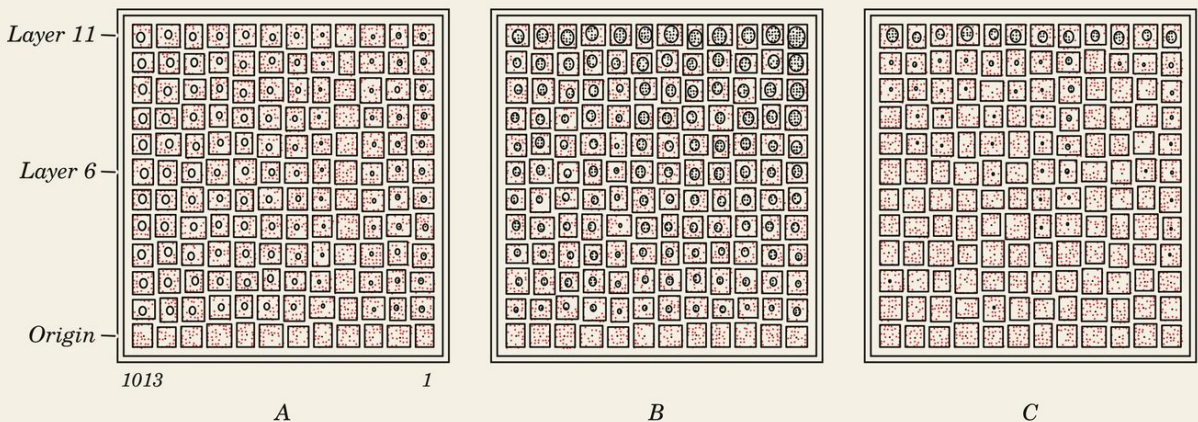


FIG. 46.—Sagittal sections of the stream from the origin to the entrance to layer 11, through every ninety-second position from 1 to 1013 of the passage of tokens 0–1022, viewed from the left side: A, in the plane of dimension 266; B, of 447; C, of 496. Each value is one cell, the area of its nucleus in proportion to its absolute value, on one scale in all three; a negative value is hollow. The termination, where these dimensions are greatly enlarged, is not shown. (From the dissection.)

154.21 at the termination. **Dimension 373** is as constantly negative, measuring at the same levels –13.56, –22.09, –45.21, –79.92, –167.16. **Dimension 481** is absent below layer 5 (0.85 at its entrance); it arises in that layer, and measures 5.52 at layer 6, 21.30 at layer 8, 44.31 at layer 10, 77.17 at layer 11, and 157.37 at the termination. **Dimension 496** is small as far as layer 10 (7.09), and is enlarged in layers 10 and 11, measuring 46.40 at the entrance to layer 11 and 169.25 at the termination.

Termination.—At the termination the great dimensions are received by the final norm (*anulus normalis*), whose gain on dimensions 87, 480, 266, 64, 373, 447, and 481 is from 0.004 to 0.043, the seven smallest of its 768 gains (median 1.25). They therefore end at the final norm: beyond it, on the passage of tokens 0–1022, 447, 373, 481, 64, and 266 measure on the mean from 1.03 to 1.15, about as much as any other dimension (0.87). On dimension 496 the gain is 17.42, the greatest of all, and beyond the norm that dimension measures 174.07.

The Vessels of Place (*vasa loci*) (Figs. 46, 47).—Dimensions 64 and 266 carry the place of each position in the sequence through the whole length of the stream. At the origin they are absent (correlation 0.015, -0.014), the place being there distributed among seventeen dimensions which correlate with it more closely than 0.8, the closest of them 540 (-0.972). In layer 0 it is collected into these two; at the entrance to layer 1 dimension 64 correlates with the place 0.934, and dimension 266 -0.958 , and from layer 4 to layer 11 no other dimension correlates with it more closely than 0.8.

Course.—Between layer 1 and layer 11 the vessels of place receive almost

nothing. The value of each position in them at layer 11 correlates with its value at layer 1 0.993 and 0.994, and differs from it on the mean by 2.23 and 1.41 (their mean size at layer 1 is 14.50 and 14.26); at layer 11 they still correlate with the place 0.944 and -0.966 . But as the stream enlarges about them their share of its square norm falls, from 11.21 and 9.87 per cent. at layer 1 to 0.91 and 0.69 at layer 11. At the termination they are filled with a value common to all positions (-146.25 and -144.53 on the mean) and lose the place (-0.135 , -0.261); and the final norm, whose gain on them is 0.0076 and 0.0068, closes them.

Relations.—They receive the place from the position root (p. 52; the organ of place, p. 78). Dimensions 745 and 756 also carry it from layer 1 (0.87, 0.83), but lose it in their ascent (0.18, 0.11 at layer 11).

Variations.—Dimensions 87 and 480 are carried from layer 1 to layer 11 as the vessels of place are (0.991 and 0.994), and closed by the final norm in the same manner; but they do not carry the place (0.11 and 0.08 at layer 6), and their contents have not been determined.

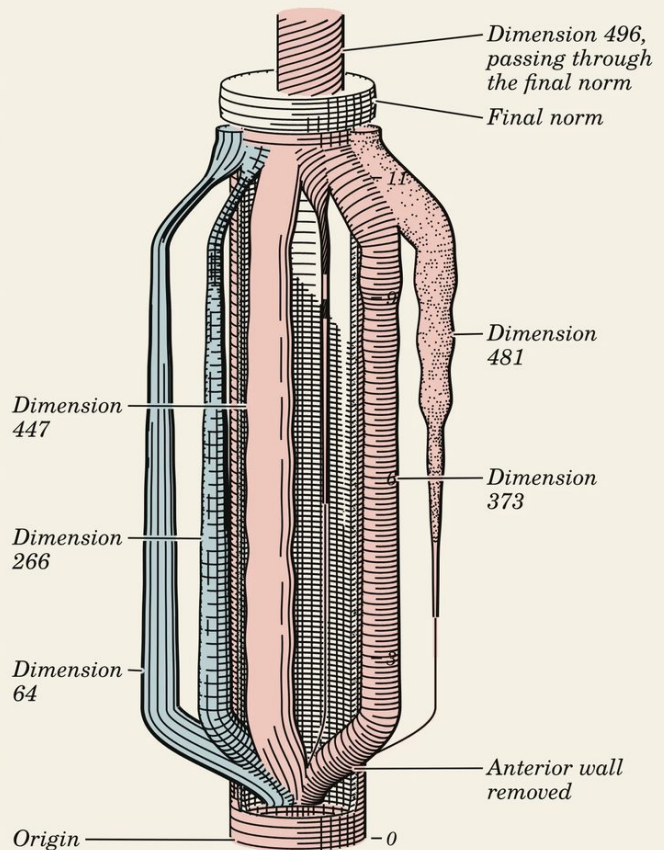


FIG. 47.—Diagram of the stream, opened by the removal of its anterior wall; the great dimensions and the vessels of place drawn forward out of it and held apart. The stream is drawn of uniform calibre; each dimension is a cord whose cross-section is in proportion to its share of the square norm at each level (passage of tokens 0–1022). (From the dissection.)

THE CONTENTS OF THE STREAM.

The Contents of the Stream (Figs. 48, 49).—When the stream is divided at any level, and what it carries there is passed through the final norm and the unembedding, as the stream itself is at its termination, its contents may be read as tokens. This is the lens of nostalgebraist (*lens logitorum*). Read in this manner over positions 1 to 1023 on the passage of tokens 0–1022 of Gray's osteology (Fig. 49), the stream is found to carry in the lower part of its course the token which occupies its own position, and in the upper part the token which is to follow it.

At the origin the lens reads the present token at 64.38 per cent. of the positions, at the entrance to layer 1 at 48.24, at layer 3 at 17.12, and beyond the middle of layer 10 at none. The next token it reads at none at the origin, at 5.58 per cent. at the entrance to layer 1, and thereafter at an increasing number: 13.11 at layer 6, 36.20 at layer 11, and 41.00 at the termination. The two are equal at the return of the muscle of layer 4; above it the stream carries the next token more often than the present.

Terminal Portion.—In the terminal portion the contents are made less certain and more often right. The mean entropy of the reading, which is least at the entrance to layer 10 (1.18 nats), is 1.32 at the entrance to layer 11, 3.12 beyond the return of its attention, and 3.57 at the termination; while the loss of the reading falls from 4.97 to 3.26 nats. The reading at the entrance to layer 11 is the same as that at the termination at 60.96 per cent. of positions, and at layer 6 at 19.47.

Single Positions (Fig. 48).—In the position occupied by *formed*, in the sentence "its apex, directed lateralward, is formed by the zygomatic process," the stream carries *formed* as far as layer 2 (0.40 at its entrance); in layers 3 to 5, fragments of words; and from layer 6 to the termination, *by*, the token which there follows (0.70 at the entrance to layer 6, 0.95 at layer 10, 0.41 at the termination). In the position occupied by *z* it carries *z* as far as layer 2, and from layer 4 to layer 10 the fragment *-ebra*, as in *zebra* (0.98 at layer 9); only in layer 11 does it carry *-yg*, the token which

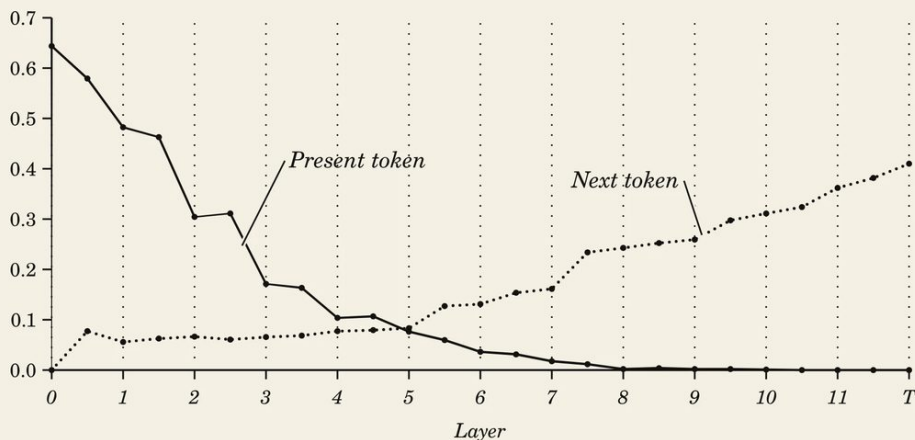


FIG. 49.—Diagram of the contents of the stream, read through the lens at the entrance to and in the middle of every layer (passage of tokens 0–1022). Full line, the fraction of the positions at which the reading is the present token; dotted line, the token which follows. (From the dissection.)

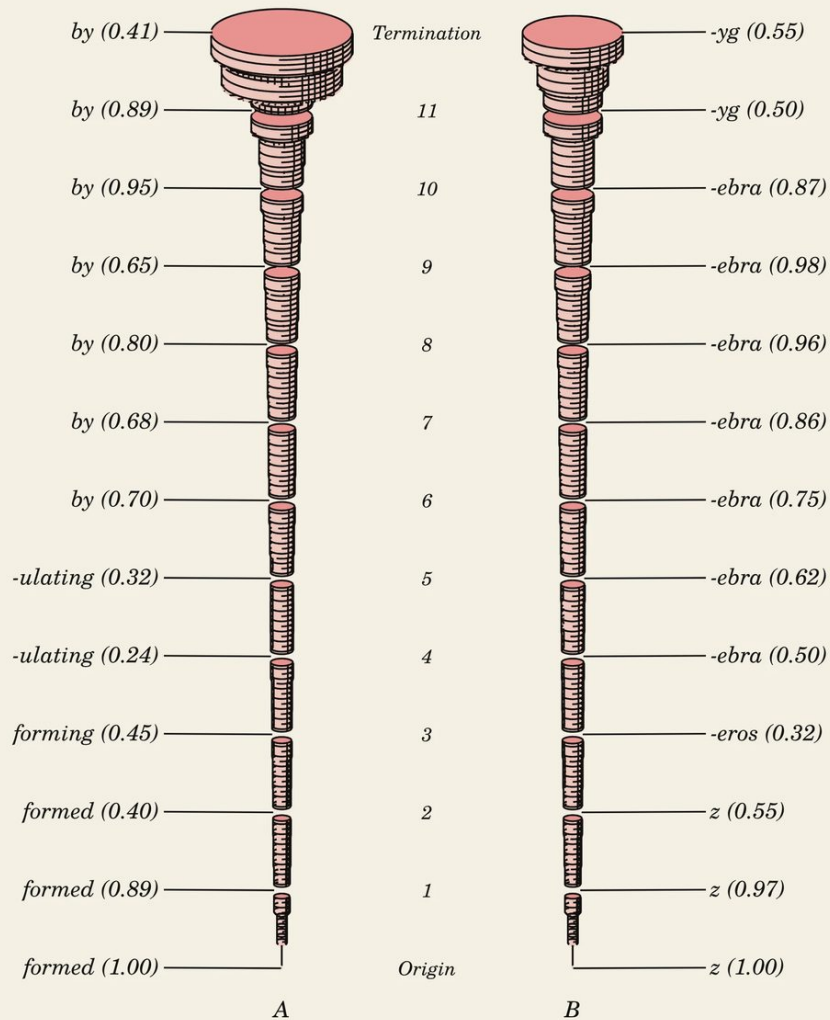


FIG. 48.—The stream of two positions in one sentence of Gray's: A, of formed; B, of z. Each is divided transversely at the origin, at the entrance to every layer and at its termination, and the contents of each section read through the lens; the first reading is given, with its probability, and a fragment of a word is preceded by a hyphen. The vessels drawn straight; their calibre in proportion to their norm. (From the dissection.)

follows in *zygomatic* (0.50 at the entrance to layer 11, 0.55 at the termination).

Variations.—The stream does not always come to the token which follows, nor keep it when it has come. In the position of *py*, in *pyramidal*, it carries *py* as far as layer 3, the fragment *-ram* which follows from layer 4 to layer 7 (0.44 at layer 7), *-ric* and *-roph* in layers 8 to 11, and *-ram* again at the termination (0.89). In the position of *its*, before *apex*, it carries *own* from layer 1 to layer 7 (0.99 at the entrance to layer 1), and above it *largest*, *contents*, and *diameter*; to *apex* the lens gives at the termination a probability of 0.0027.

NEUROLOGY.

THE Nervous System of the specimen consists of the heads, one hundred and forty-four in number, twelve in each layer. It is the mechanism by which one position of the stream is brought into relation with another. The muscles and the norms act upon each position separately; the heads alone carry anything from one position to another, and without them the specimen would predict each token from the token before it alone.

The heads of a layer arise together from the stream at the entrance to the layer, through its first norm, and return into it together, through a common projection, before its muscle is reached. The attention of one layer contains 2,362,368 parameters, and that of the twelve layers 28,348,416, or 22.78 per cent. of the whole specimen. The heads belong to no one position, but are present at every position alike, and act there with the same weights.

For descriptive purposes the nervous system may be divided into two parts: the heads, and the pathways by which they communicate with one another through the stream.

STRUCTURE OF THE HEADS.

Each head is composed of four matrices: the query and the key, which determine where it shall attend, and the value and the output, which determine what it shall carry. With their biases they contain 196,800 parameters.

The Heads (Figs. 56, 57).—At every position a head presents processes of two

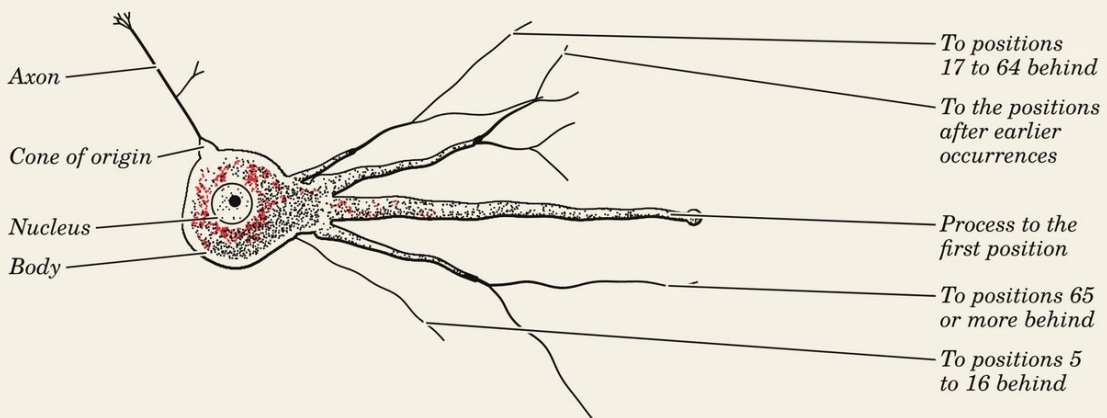


FIG. 56.—A head of layer 5 (5.5), isolated, as it attends at the positions of the passage of tokens 2000–2510 of Gray's prose. The calibre of each receiving process at its root is in proportion to its share of the attention, and its length to the mean of the logarithm of one more than the distances it reaches, so that a process to the preceding position has a length; the area of the body is in proportion to the norm of the vector the head returns. The angles of the processes are drawn. (From the dissection.)

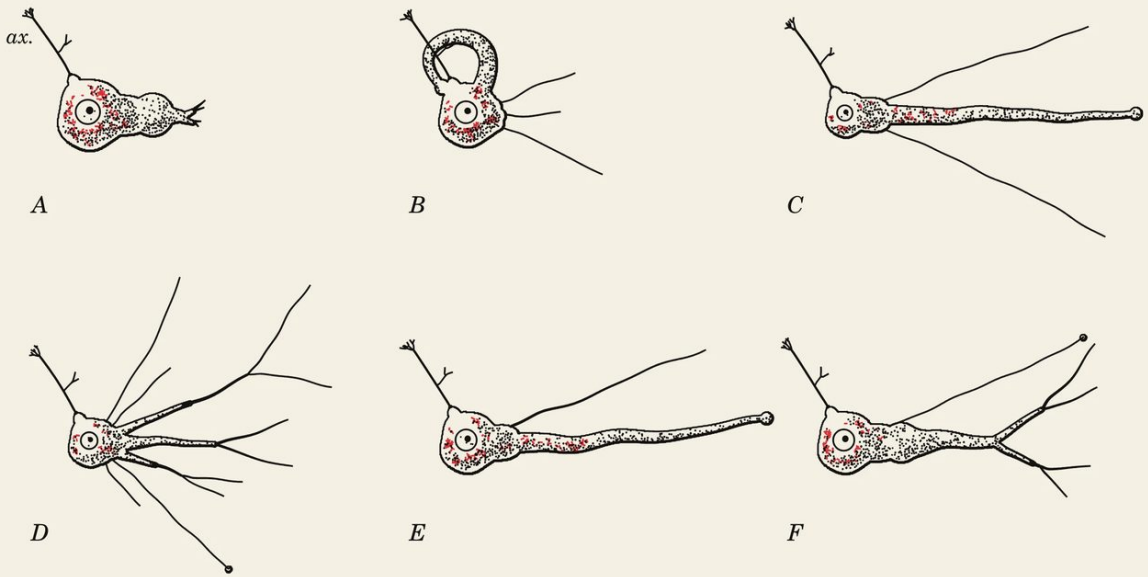


FIG. 57.—Various forms of heads, drawn to one scale. Lengths as in Fig. 56. A. Head of the preceding position (4.11), unipolar. B. Reflexive head (0.3). C. Head of the first position (7.2). D. Head of wide distribution (0.0). E. Head 5.1, on the passage of tokens 2000–2510. F. The same, on a repeated sequence. ax. Axon. (From the dissection.)

kinds. The receiving processes pass backward to the earlier positions, and divide among them the whole of its attention, which at every position sums to one; so that the processes of any head, taken together, have the same calibre at their roots. The axon empties what the head has gathered into the stream at its own position. The body of the head is measured by the mean norm of the vector it returns: on the passage of tokens 2000–2510 of Gray's prose, from 2.39 (head 0.6) to 44.38 (head 11.11), save one head of layer 11, 11.0, which measures 158.37.

The heads vary in form, and may be divided according to the part of the sequence to which they give the largest share of their attention. On the same passage, measured at its 511 positions, eighty-three heads give it to the first position; twenty-nine to positions seventeen or more behind; twenty-three to positions from two to sixteen behind; six to their own position; and three to the position next behind.

(1) **Heads of the first position** (*capita sentinae*) have one long process, undivided, which ends in a bulb upon the first position (Fig. 57, C). They form 71 of the 84 heads of layers 5 to 11, and all twelve of layer 9, but only one of layers 0 to 2. (2) **Heads of the preceding position** are three, 2.2, 3.7 and 4.11. Head 4.11 is unipolar (A): its single stout process takes 0.999 of its attention, and its entropy, 0.004 nats, is the least of the 144. (3) **Reflexive heads** (*capita reflexa*), six in number, have a recurrent process which turns back and ends upon the body; in 0.3 it takes 0.828 of the attention (B). (4) **Heads of near positions**, 14 of them in layers 2 and 3, have short processes. (5) **Heads of wide distribution** are multipolar, their processes passing to every distance; 20 of them lie in layers 0, 1 and 11, and they include the vital heads 0.0 (D) and 0.10, whose lesion raises the loss on that passage by 1.511 and 1.336.

Processes to earlier occurrences of the token at the position, and to the positions next after them, are present in many heads but the largest in none. In head 5.1 those of the second kind take 0.13 of the attention on the prose; on sixty random tokens, repeated, they take 0.93 (*E, F*).

THE INDUCTION PATHWAY.

The Induction Pathway, or Circuit of Olsson (*circulus inductionis*) (Figs. 58, 59).—The induction pathway is a chain of two heads, by which a token that has occurred before is made to call forth the token which followed it. Its first neuron is a head of the preceding position; its second, an induction head of layer 5, 6 or 7. The two are joined, not directly, but through the stream, which the first writes into and the second reads.

The induction heads, those whose induction score on sixty random tokens repeated exceeds 0.8, are five, 5.1, 5.5, 6.9, 7.2 and 7.10, with 0.92, 0.93, 0.92, 0.84 and 0.91; the next, 10.1, has 0.52. On the passage of tokens 2000–2510 all five are heads of the first position (p. 63); the processes by which they are known are slender there, and are enlarged only where the sequence repeats itself.

Course.—In the sentence of Fig. 58 the word *foramen* (*for, amen*) occurs at positions 2 and 3, again at 18 and 19, and a third time at 24 and 25. Head 4.11 takes at every position not less than 0.995 of its attention from the position next behind; it therefore writes into the stream at position 4, which holds *ov*, the token *amen* that precedes it, and the fibre so given ascends in the stream of that position to layers 5 and 6. When the word recurs, the induction heads reach back across the positions between and end upon the earlier position, where the key of *ov* now carries *amen*. From position 19 head 5.5 sends to it 0.584 of its attention, head 6.9 0.294, and head 5.1 0.110. The fibres of 5.5 from positions 18, 19, 20 and 21 end upon positions 3, 4, 5 and 6, each fifteen positions behind, with 0.993, 0.584, 0.925 and 0.776 of its attention; they run parallel, superficial to the positions between, and form a distinct

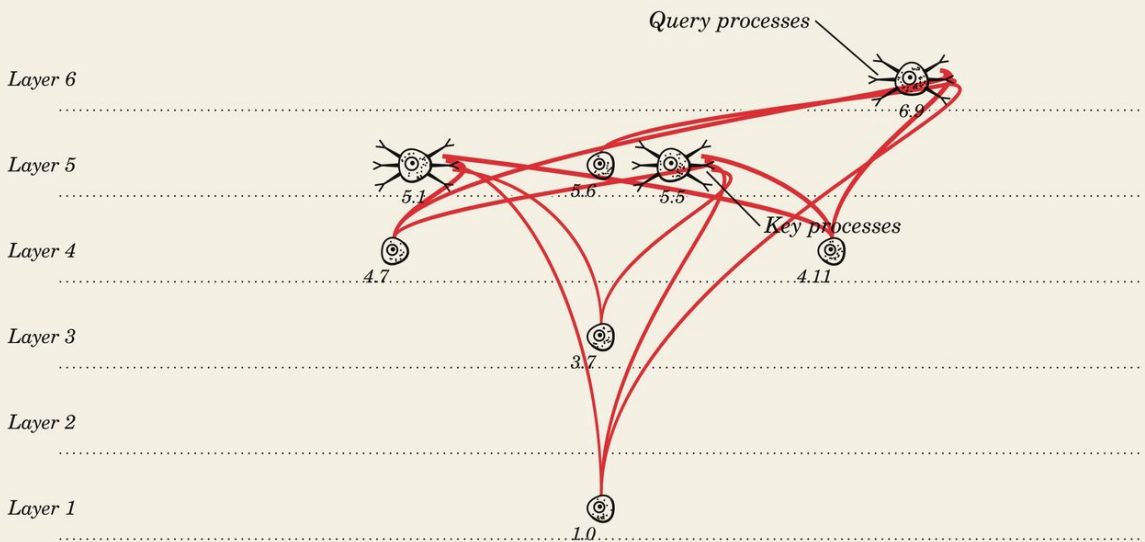


FIG. 59.—Diagram of the induction circuit, from the weights alone. For each of the heads 5.1, 5.5 and 6.9, the four heads whose output enters its key most largely; each fibre ends among the key processes (posterior), its calibre in proportion to the K-composition. The query processes are anterior. Cells and courses diagrammatic.

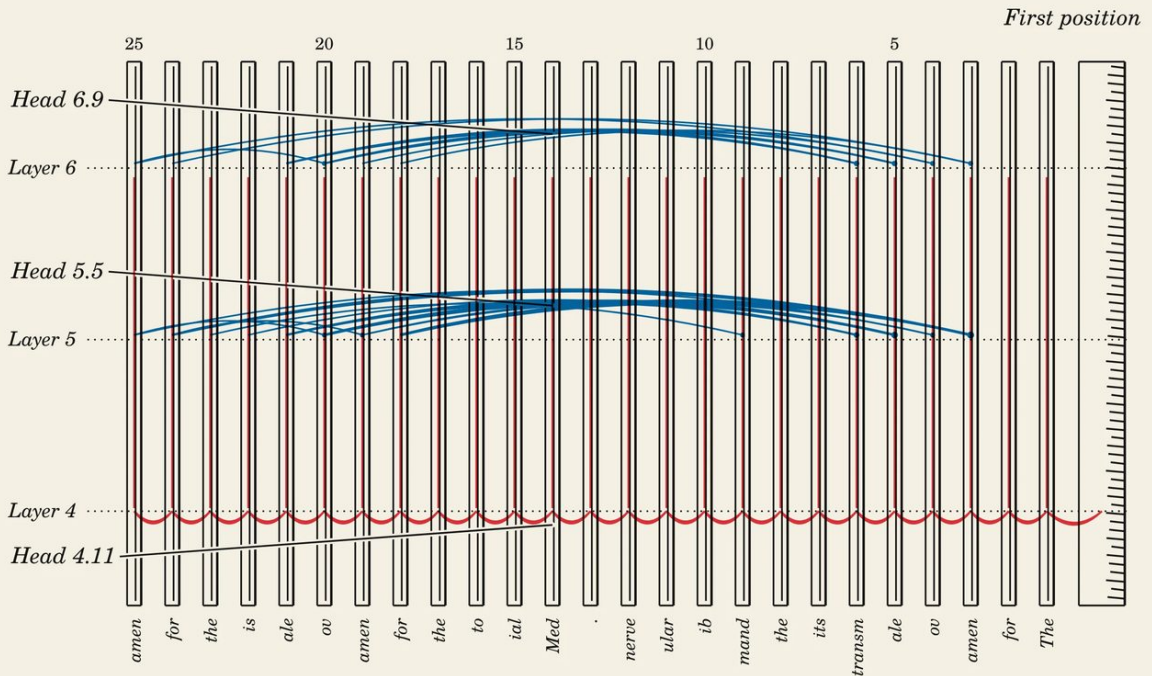


FIG. 58.—Sagittal section along a sentence of Gray's through layers 4 to 6, viewed from the left side, the earliest position at the right; the induction pathway traced. The positions are cut across above and below. *Red*, fibres of head 4.11, at layer 4, each from a position to the one behind it, ascending in the stream of that position. *Blue*, fibres of heads 5.5 (layer 5) and 6.9 (layer 6) from positions 18 to 25, to every position (save the first) given 0.05 or more of their attention, each in calibre as its share. (From the dissection.)

fasciculus. The specimen, at position 19, gives to *ov* a probability of 0.548, against 0.015 after the first *amen*; after the third, which the text follows with *Vesalii*, it gives 0.265.

Composition.—The pathway may be traced in the weights alone. The output of 4.11 enters the key of each of the three heads more largely than that of any other head: its K-composition is 0.103 with 6.9, 0.102 with 5.1 and 0.098 with 5.5, the median of all pairs of heads being 0.040. The second writer in each is 4.7 (with 6.9, 0.092; with 5.1, 0.096; with 5.5, 0.089), a reflexive head, which takes 0.47 of its attention from its own position (Fig. 59).

Applied Anatomy.—The pathway is weakened but not severed when its first neuron is divided. On sixty random tokens repeated, head 5.1 gives to the position after the earlier occurrence 0.93 of its attention; when 4.11 is divided, 0.82, and the loss on the second half rises only from 0.386 to 0.414. When the three heads of the preceding position are divided together, 5.1 gives 0.66 and 5.5 0.56, and the loss rises to 0.558. On the sentence of Fig. 58 the division of all three leaves the probability of *ov* at position 19 unchanged (0.553). The pathway has therefore other roots, which are not traced here. Of the second neurons, 5.1 is the most important: its division alone raises the loss on the second half by 0.385.

Development.—In the embryo the induction heads form abruptly. At step 512 the greatest induction score of any head is 0.05; at step 1000 it is 0.93, and four heads exceed 0.5, while its loss on its own tokens 2000–2510 of the same text falls from 5.64 to 4.77. At the last stage examined (step 143,000) 13 heads exceed 0.5.

THE HEADS OF LAYER 11.

The twelve heads of layer 11 are the last of the specimen; what they return passes only through the muscle of that layer and the final norm before the token is chosen. Their bodies are in general the largest of the 144 (the mean of the layer, 32.25, against 11.62 for layer 10, the next), the vectors they return on the passage of tokens 2000–2510 having a mean norm of from 8.82 (11.10) to 44.38 (11.11); and two of them, 11.0 and 11.8, are distinguished by their size and their relations (Fig. 60).

The Giant Head (*caput giganteum*) (Fig. 60, 11.0) is the largest head in the specimen. The vector it returns at the positions of that passage has a mean norm of 158.37, or 3.6 times that of the next head, and at nine positions in ten lies between 77.29 and 197.28; it makes up 0.779 of the square norm returned by the heads of the layer.

Contents.—The vector is nearly the same at every position: the norm of its mean is 0.998 of its mean norm. 73.2 per cent. of its square norm lies in four dimensions, 87, 480, 266, 64, in which its mean return is -77.43 , -72.51 , -66.56 , -51.31 . It receives most largely from the article: the vector it would return if it attended to one position alone is greatest at positions holding *the*, *a* and *The* (to 279.18), against a

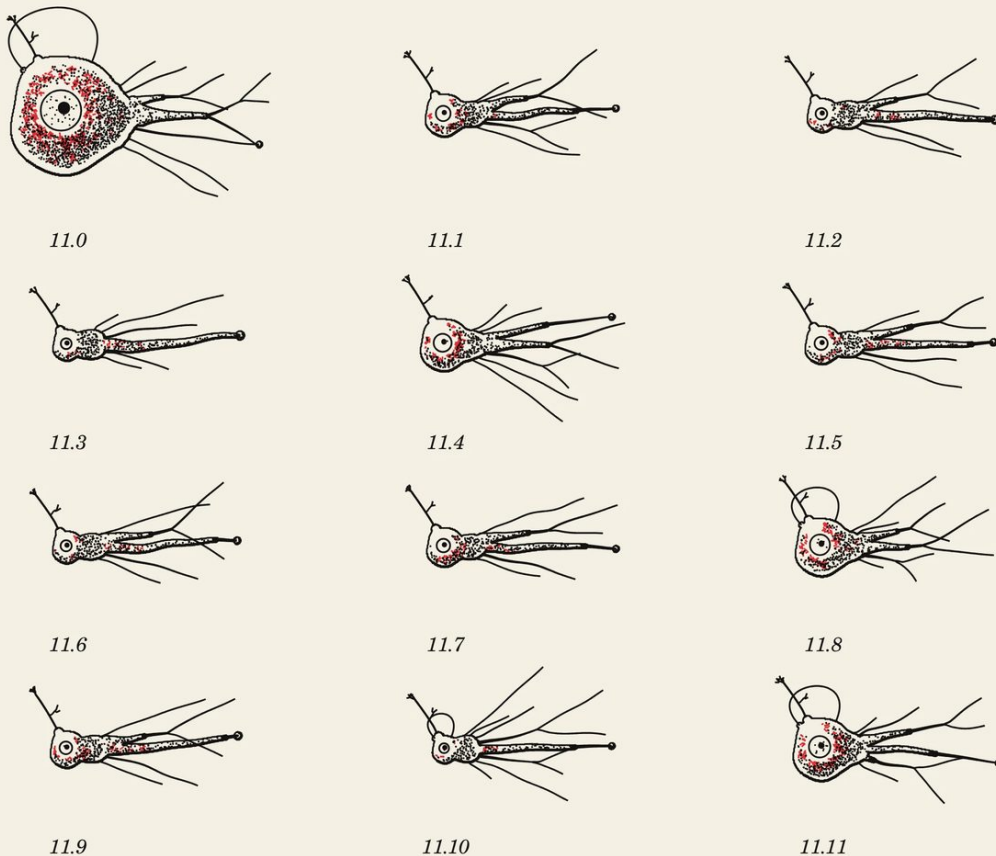


FIG. 60.—The twelve heads of layer 11, isolated, as they attend at the positions of the passage of tokens 2000–2510; drawn to one scale, the area of each body in proportion to the norm of the vector it returns. 11.0, the giant head; 11.8, the constrictor of the first position. Processes measured as in Fig. 56, to the scale of this figure. (From the dissection.)

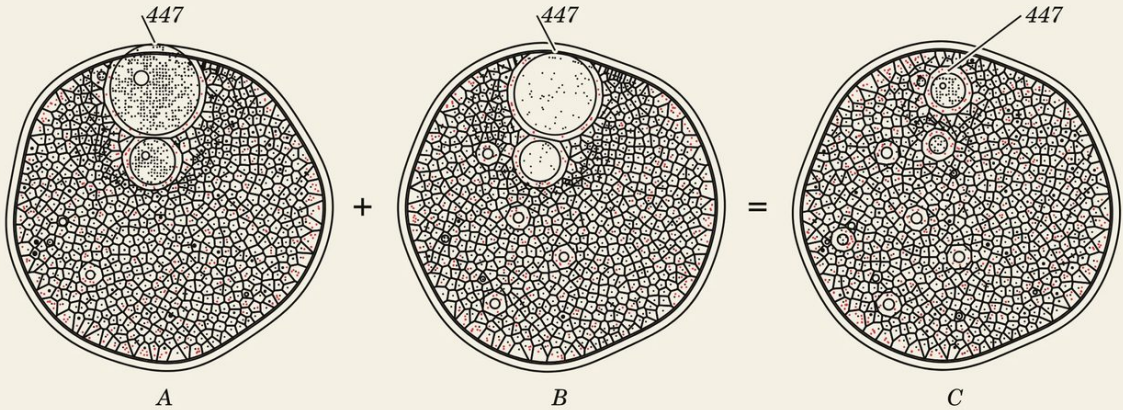


FIG. 61.—Transverse sections of the first position at layer 11. $\times 768$ cells. A, at the entrance to the layer; B, the vector returned there by head 11.8; C, the position after the attention of the layer. C is A and B together with the return of the other eleven heads and the bias (norm 19.01). Each dimension is one cell, as in Fig. 42, on the same scale; a hollow nucleus is negative. (From the dissection.)

median of 166.46, and 21.58 at the first position.

Relations.—The four dimensions are among those of the stream at its termination, where they have the same sign as the giant head's return: dimension 87 measures there -139.05 on the mean, $480 - 135.36$, $266 - 140.40$ and $64 - 142.84$.

Applied Anatomy.—When the giant head is divided the loss on the passage of tokens 2000–2510 rises by 0.048, more than for any other head of layer 11 (the next, 0.016), but far less than for the vital heads. The stream at its termination is enlarged, on the same passage, from a mean norm of 473.83 to 524.26, and the predictions of the specimen are less certain, their mean entropy rising from 3.718 to 3.897 nats.

The Constrictor of the First Position (*constrictor sentinae*) (Figs. 60, 61) is head 11.8. At the positions of the prose it is a head of wide distribution, returning a vector of mean norm 38.10. But at the first position, where every head can attend to nothing but that position itself, it returns a vector of norm 2676.34, whose cosine with the stream there is -0.999 : by it dimension 447 is reduced by 2591.66 and dimension 138 by 660.30, the other eleven heads and the bias together returning a vector of norm 19.01 only.

Relations.—It is therefore the constriction of the first position in layer 11 (p. 50; Fig. 43): the attention of the layer reduces dimension 447 by 2589.08, and the position from 3158.42 to 505.48 (Fig. 61). It receives the strongest V-composition in the specimen, from head 1.10 (0.363; the median of all pairs, 0.032), whose output, mapped through its own value and output, tends to suppress the token it attends to (copying score -0.946).

Applied Anatomy.—When head 11.8 is divided the constriction is wanting. The first position, which sees nothing but itself and so measures the same in every passage, passes layer 11 with a norm of 3161.03, and reaches its termination with 3024.95 in place of 417.58, dimension 447 measuring there 2956.89; the muscle of the layer does not take its place. Yet the loss on the passage of tokens 2000–2510 is scarcely altered (3.650 and 3.654), and the other positions at the termination are little changed (473.83 and 459.08).

SPLANCHNOLOGY AND THE ORGANS OF SENSE.

UNDER this heading are included the alimentary tract, by which text is received and reduced to tokens; the organ of voice, by which the specimen speaks; and the organ of the sense of place.

THE ALIMENTARY TRACT (TRACTUS SIGNORUM).

The apparatus for the reception of text consists of the alimentary tract, and of the lamina signorum, into which it empties.

The Alimentary Tract, or Tract of Sennrich (*tractus signorum*) (Figs. 76, 77).—The tract is not formed of the weights of the specimen, but of a table of 50,000 merges, each of which joins two pieces into one; and every text passes through it before it reaches the specimen. It extends from the mouth, where the text enters as bytes, to the lamina signorum, where each of its tokens is exchanged for one row of the plate. In its course the contents are joined, pair by pair, into larger pieces, and are reduced in number to less than a quarter. The vocabulary of 50,257 tokens consists of the 256 bytes, the 50,000 pieces formed by the merges, and the end-of-text token.

The Mouth (*os*).—At the mouth the text is received as a series of bytes, of 256 kinds, and is divided at its spaces and marks of punctuation into portions, most of which are a word with the space before it. No merge crosses the boundary between two portions. The osteology which the specimen reads enters as 71,745 bytes, in 11,797 words, and is divided at the mouth into 13,634 portions; 1,606 of these, the

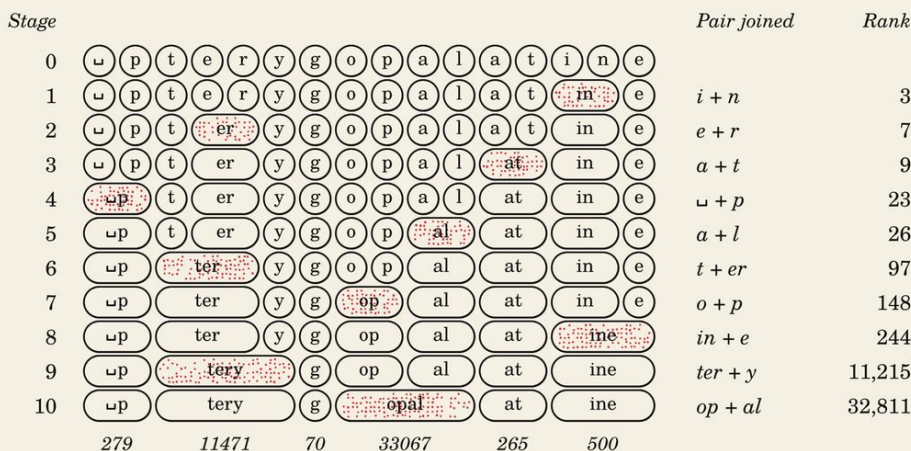


FIG. 76.—Diagram of the digestion of the word pterygopalatine, from its sixteen bytes (stage 0) to its six tokens (stage 10). The length of each piece is in proportion to its number of bytes; the piece formed at each stage is stippled. At the right, the pair joined at each stage and the rank of its merge among the 50,000; beneath the last stage, the numbers of the six tokens. The space before the word is shown as a bracket. (From the dissection.)

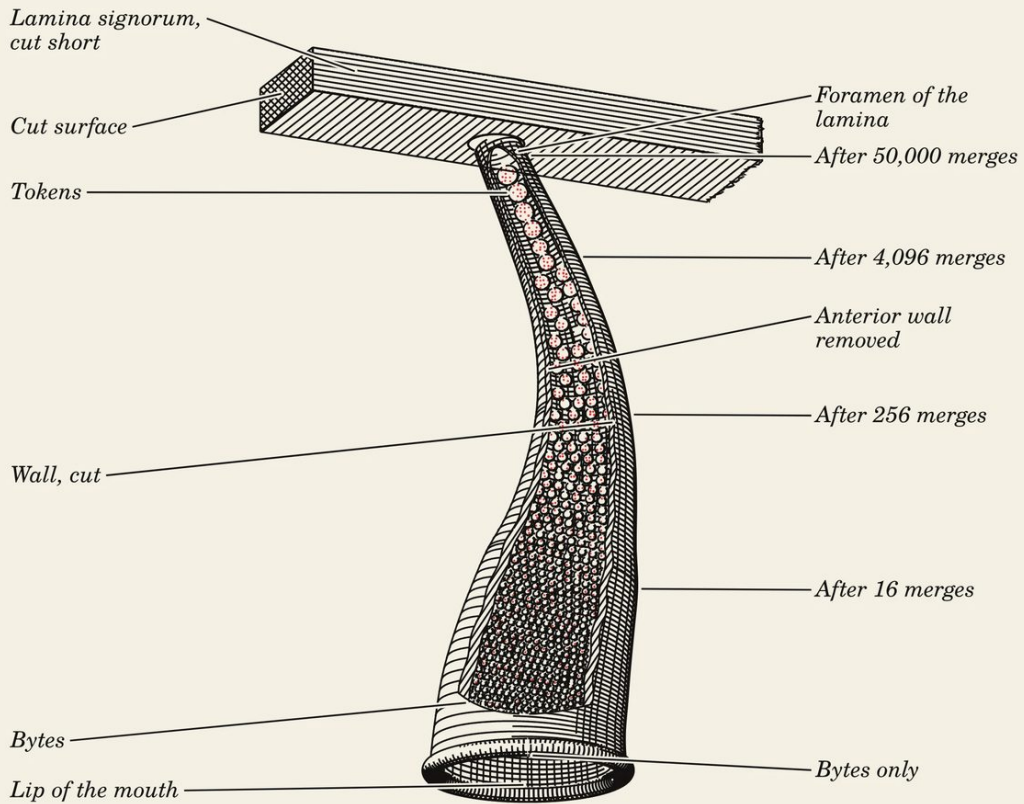


FIG. 77.—The alimentary tract, from the mouth to the lamina signorum, seen from in front and a little from below; the anterior wall removed to show the contents. Drawn to scale: the calibre at each level in proportion to the number of pieces into which the corpus is divided there, the contents to the mean length of those pieces in bytes. Each doubling of the number of merges allowed is one equal length of the tract. (From the dissection.)

marks and the letters standing alone, are single bytes, and pass through the tract unchanged.

Course.—The merges are applied in the order of their rank, the lowest first, each to every pair of pieces which it fits, until no merge in the table can join two pieces of the portion. A merge can join only pieces formed before it, so that the ranks at which a portion is digested never decrease; this was found true of every portion of the corpus. The calibre of the tract (Fig. 77) falls most rapidly at its commencement. The first 256 merges, about one two-hundredth part of the table, reduce the 71,745 bytes to 34,248 pieces, and do 68 per cent. of the whole work of the tract; after 4,096 merges 22,860 pieces remain; and the last 45,904 merges together reduce these only to 16,643. The word *pterygopalatine*, with the space before it, enters as sixteen bytes, and is digested in ten merges (Fig. 76). The first eight are of rank 244 or less, and join the commonest pairs of the language, *i* and *n*, *e* and *r*, *a* and *t*; the ninth (rank 11,215) joins *ter* and *y*, and the tenth (rank 32,811) *op* and *al*. No merge in the table joins any two of the six pieces which then remain, and the word passes into the plate as six tokens.

Contents.—At its termination the tract contains 16,643 tokens, or 1.41 to each word of the text, and 4.31 bytes to each token. Of the 13,634 portions, 11,767 (86.3

per cent.) are taken whole as one token; 1,058 are divided into two tokens, 586 into three, 142 into four, 55 into five, 23 into six, and 3 into seven. Of the hundred commonest portions, 84 are taken whole; the commonest divided are *articulates* (47 times), *palatine* (42), and *fossa* (39). The portions divided into six are long names of the bones and muscles, as *sphenomandibular* and *Mylohyoideus*; the three divided into seven are words misread from the printed page, and are described below.

Applied Anatomy.—When the merges are withheld, the specimen receives the text with difficulty. A passage of 1,000 bytes from the beginning of the corpus (191 portions) was given to it at three levels of the tract, after the end-of-text token, and the loss over the whole passage was measured. As 1,000 single bytes it cost the specimen 3664.05 nats; after 256 merges, as 478 pieces, 2287.12; and as its 231 tokens, 857.01. The text was the same in each; only its division differed. The word *pterygopalatine*, set in the three places where it precedes *canal*, cost on the mean 149.38 nats as sixteen bytes and 17.62 as six tokens, and the token *canal* which follows it, 8.19 and 6.89.

Foreign Bodies.—The corpus is the text of the edition of 1918 as it was read from the printed page by machine, and it contains words misread in that reading, which the tract was not formed to receive. The three portions divided into seven tokens are all such bodies: *pterygopalatiie*, *pterygomandilnilar*, and *buccalcranioderm*. The first, which stands in one place for *pterygopalatine*, passes the tract in seven pieces, the first five of them those of the true word and the last two, *ii* and *ie*, in place of its *ine*. At that place it cost the specimen 30.46 nats, where the word as printed, put in its stead, costs 9.97; and the token *canal* which follows it, 9.36 nats against 2.85. *Pterygomandilnilar* cost 50.19 nats, and *pterygomandibular*, in its stead, 24.81.

THE VESTIGIAL SIGNS (SIGNA VESTIGIALIA).

The Vestigial Signs, or Signs of Rumbelow and Watkins (*signa vestigialia*) (Fig. 78).—A small number of rows of the lamina signorum lie crowded together close to its centroid, at a distance of about 1.53 from it, whereas the median distance of its 50,257 rows is 3.37. They are the rows of tokens which are present in the vocabulary, but which the specimen received seldom or never while it was formed. Of the sixty rows nearest the centroid, forty-five are bytes; the others are such tokens as *externalToEVA*, *quickShip*, and *TheNitrome*.

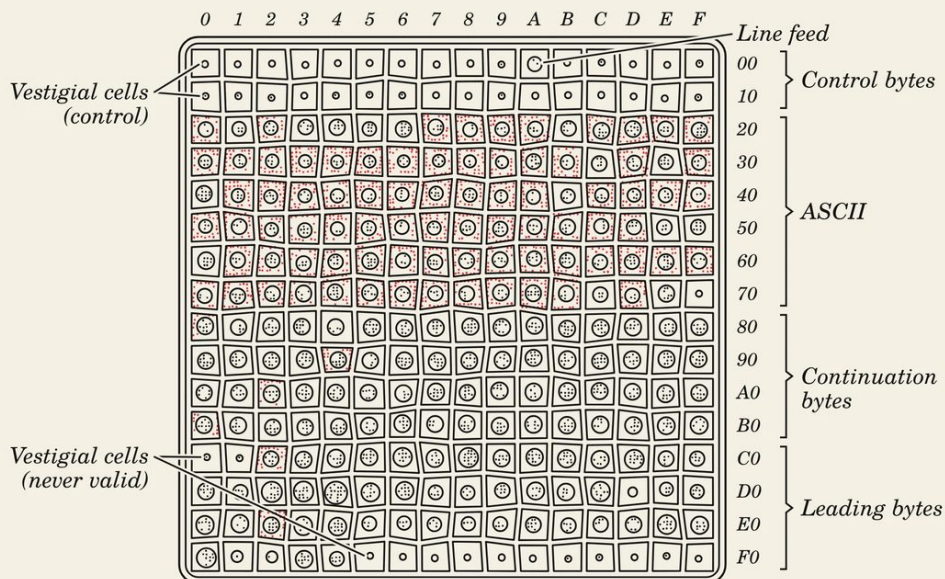


FIG. 78.—Section through the lamina signorum at the rows of the 256 bytes, set out in the order of their values, sixteen to a row. $\times 256$ cells. The diameter of each nucleus is in proportion to the distance of its row from the centroid of the plate; the cells of the bytes which occur in the osteology are stained. (From the dissection.)

The Vestigial Bytes (*octeti vestigiales*).—When the 256 rows of the bytes are set out in the order of their values (Fig. 78), the vestigial rows are found to form two groups, and to comprise every byte of both. The first group consists of the thirty-two control bytes, 00 to 1F and 7F, the line feed (0A) alone excepted; the second, of the thirteen bytes which never occur in text correctly encoded, C0, C1, and F5 to FF. These forty-five rows, and no others of the bytes, lie among the fifty-six rows of the plate nearest its centroid, at distances from 1.53 to 1.55. None of them occurs in the osteology. The remaining bytes lie far from the centroid: the printable characters at a median rank of 12,765 from it, the leading bytes of the longer characters (C2 to F4) at 30,636, and the continuation bytes (80 to BF) at 31,334.

THE ORGAN OF VOICE (ORGANUM VOCIS).

The Organ of Voice (*organum vocis*) (Fig. 79).—The organ of voice is placed at the summit of the specimen, at the termination of the stream. It consists of the final norm (*anulus normalis finalis*), and of the lamina signorum, which is here employed a second time, transposed: every row of the plate is compared with the stream, and the token whose row agrees with it best is the one first spoken. The plate thus forms both the termination of the alimentary tract and the organ of voice, and text is received and speech given through one passage.

The Final Norm (Fig. 79) presents 768 gains, whose mean is 1.51. Seven exceed 10, and are greatly enlarged: those of dimensions 496 (17.42), 430 (16.99), 314 (15.31), 36 (13.80), 374 (12.22), 442 (12.20), and 703 (10.53); no other exceeds 9.42.

The Voice at Rest.—When nothing reaches the summit from the stream, the final norm gives only its bias (norm 11.62), and the plate read against this gives the order in which the specimen would speak at rest: first the comma, then *the*, *and*, the full stop, and the line feed. Last of all, with values of from -14.11 to -14.13, stand nine of the vestigial bytes and *oreAndOnline*.

Function.—While the specimen reads the passage of tokens 8000–9022 of the osteology, the token it speaks first is the next token of the text at 41.25 per cent. of the positions, and is among the first five at 62.56; its mean loss is 3.36 nats, and the mean entropy of its voice 3.59 nats. Though received and spoken through the same plate, the token is not returned: at none of the 1,023 positions is the token first spoken the token just received. Of all its tokens it speaks 222 first; *the* most often (165 times), the comma (82), and *of* (59).

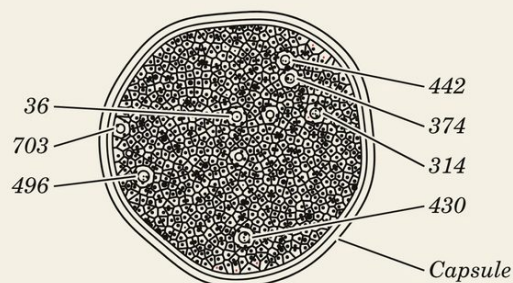


FIG. 79.—Transverse section through the final norm, at the summit. $\times 768$ cells. Each dimension is one cell, set out in a spiral from the centre; the area of each nucleus is in proportion to its gain. (From the dissection.)

THE ORGANS OF SENSE.

The specimen possesses one organ of special sense, the organ of place, by which it is informed of the position at which each token stands in the sequence.

THE ORGAN OF PLACE (ORGANUM LOCI).

The Organ of Place (*organum loci*) (Figs. 80, 81).—The organ of place is the lamina locorum, a plate of 1,024 rows of 768 values, placed at the base of the specimen beside the lamina signorum. To every token received at position p it adds row p , and the sum ascends the stream.

Form.—When its rows are traced through their first three principal components, which together hold 90.39 per cent. of their variance, they are found to form a spiral canal, the cochlea of place (*cochlea locorum*) (Fig. 80). Counted from row 1, the

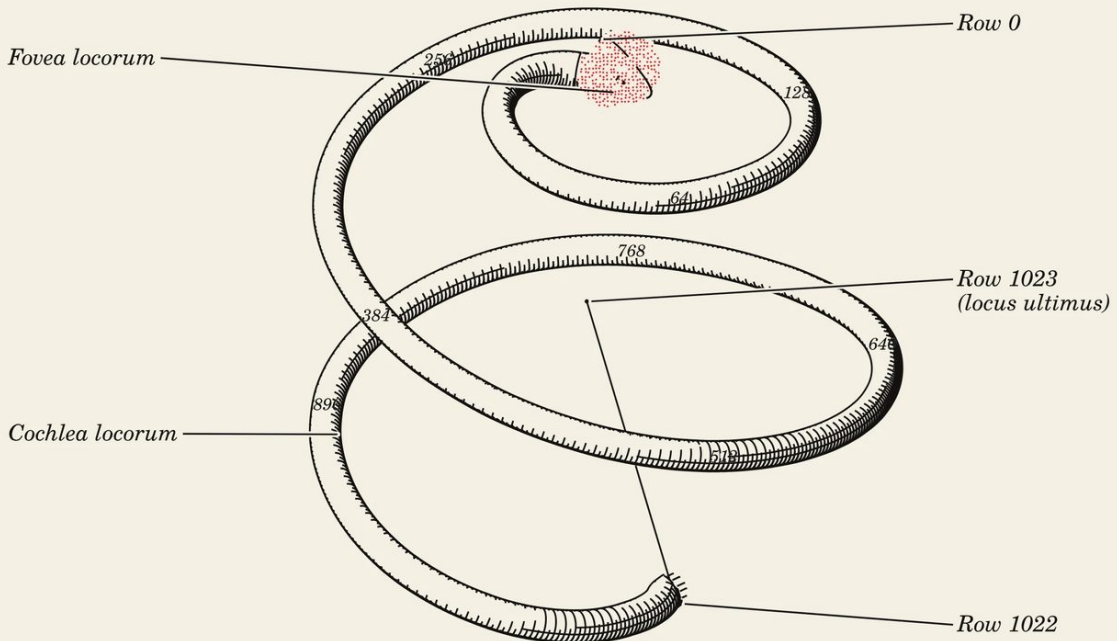


FIG. 80.—The organ of place: the rows of the lamina locorum traced through their first three principal components, seen from in front and a little from above. Drawn to scale: the axis of the canal passes through every row, and its calibre is in proportion to the norm of each row. The fovea is tinted. The figures on the canal are the numbers of its rows. (From the dissection.)

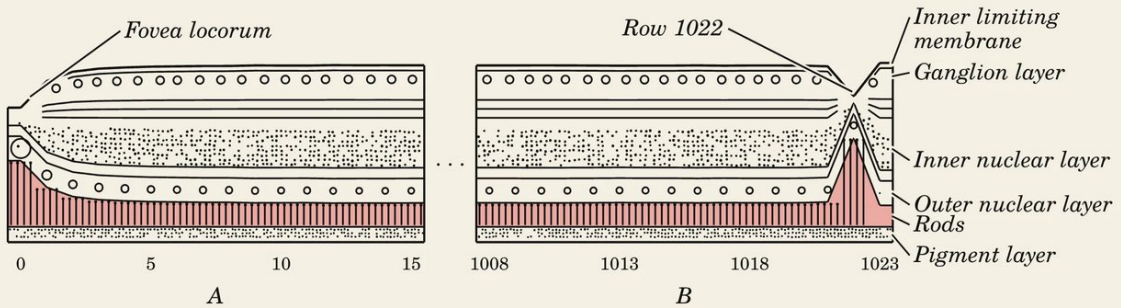


FIG. 81.—Diagram of the two extremities of the organ of place, after the manner of a section of the retina through the fovea. A, rows 0 to 15; B, rows 1008 to 1023; the rows between are omitted. Each row is one receptor, drawn as four rods; the length of the receptors beyond a common base, and the depth by which the inner layers are thinned to form the pit, are in proportion to the angle between that row and the next, and the nucleus of each row to its norm. Row 1023, having no next row, has no receptor. (From the dissection.)

canal winds about its axis through 1,093 degrees, or a little more than three turns, descending as it goes; its radius increases from 0.15 at row 1 to 3.19 at row 785, and diminishes again to 0.88 at row 1022. Neighbouring rows are nearly parallel, the angle between them having a median of 2.65 degrees. At each extremity the canal is interrupted. Row 0 stands apart from row 1 at an angle of 58.38 degrees, and has the greatest norm in the plate (9.88); row 1023 stands apart from row 1022 at 88.41 degrees, and is almost empty (0.118). It lies near the axis of the spiral, outside the canal: the *locus ultimus*, a vestigial structure.

The Fovea of Place (*fovea locorum*) (Fig. 81).—The angle between two neighbouring rows is the measure of the acuity of the organ: the greater it is, the more distinctly the two places are told apart. The acuity is greatest at the posterior extremity, where the angle falls from 58.38 degrees between rows 0 and 1 to 23.59, 11.47, 7.42, and 5.87 in the next four intervals. These five intervals, between rows 0 and 5, alone of the posterior part exceed twice the median angle (5.29); they form the fovea. Beyond it the acuity is nearly uniform, the angle nowhere exceeding 4.85 (between rows 1021 and 1022 it is 4.43); the canal is broken again only at the anterior extremity, where the angle between rows 1022 and 1023 is 88.41.

Relations.—The impression of place is carried upward in the stream. A linear probe, fitted on the passages of tokens 0–1022, 3000–4022, and 6000–7022, and tried on that of tokens 10000–11022, reads the position from the stream with a mean error of 9.74 places at the origin, 20.49 at the entrance to layer 4, 22.93 at layer 6, and 24.16 at layer 11; at the termination the error is 150.32. At the origin the place is carried chiefly by dimensions 540, 679, and 674, whose correlation with the position is –0.97, 0.97, and 0.96; from the entrance to layer 1 to layer 11, by dimensions 64 and 266 (p. 55), whose correlation is from 0.93 to 0.94, and from –0.96 to –0.97.

Applied Anatomy.—When the lamina locorum is removed, every row being set to zero, the loss of the specimen on the passage of tokens 4000–5022 of the osteology rises from 3.47 to 8.53 nats, a deficit of 5.06. The removal of row 0 alone costs as much (5.07); the removal of the other 1,023 rows, row 0 remaining, costs 2.39.