

TRƯỜNG ĐẠI HỌC ĐÀ LẠT



GIÁO TRÌNH

CƠ SỞ NGÔN NGỮ HỌC

DƯƠNG HỮU BIÊN

KHOA NGỮ VĂN

MỤC LỤC

PHẦN I : BẢN CHẤT XÃ HỘI CỦA NGÔN NGỮ	7
CHƯƠNG I : NGÔN NGỮ VÀ NGÔN NGỮ HỌC	7
I. NGÔN NGỮ.....	7
1. Giới thiệu.....	7
2. Ngôn ngữ học.....	7
3. Các thành phần của ngôn ngữ.....	7
4. Việc thụ đắc ngôn ngữ.....	9
5. Hiện tượng song ngữ và đa ngữ.....	10
6. Những sự đa dạng ngôn ngữ.....	10
7. Các ngôn ngữ của thế giới.....	12
8. Ngôn ngữ phi lời nói.....	16
II. NGÔN NGỮ HỌC	17
1. Giới thiệu.....	17
2. Ngôn ngữ học miêu tả và so sánh.....	17
3. Các lĩnh vực của ngôn ngữ học.....	19
4. Lịch sử của ngôn ngữ học.....	21
CHƯƠNG II : NGUỒN GỐC HÌNH THÀNH VÀ SỰ TIẾN HÓA CỦA NGÔN NGỮ.....	24
I. MỞ ĐẦU.....	24
II. CÁC QUAN NIỆM TRƯỚC MARX VỀ NGUỒN GỐC CỦA NGÔN NGỮ.....	24
1. Nguồn gốc siêu nhiên.....	24
2. Nguồn gốc âm thanh tự nhiên.....	25
3. Nguồn gốc kết hợp miệng - điệu bộ.....	25
4. Sự thích nghi sinh-lý học.....	26
5. Tiếng nói và chữ viết.....	27
III. QUAN ĐIỂM MÁC – XÍT VỀ NGUỒN GỐC NGÔN NGỮ.....	27
IV. SỰ TIẾN HOÁ CỦA NGÔN NGỮ.....	28
1. Các quy luật của sự tiến hóa ngôn ngữ.....	28
2. Các ngôn ngữ biến đổi như thế nào?.....	30
V. KẾT LUẬN.....	32
CHƯƠNG III : BẢN CHẤT, CHỨC NĂNG VÀ THUỘC TÍNH CỦA NGÔN NGỮ.....	32
I. BẢN CHẤT XÃ HỘI CỦA NGÔN NGỮ.....	32
1. Ngôn ngữ là sản phẩm của con người, của xã hội.....	32
2. Ngôn ngữ là một hiện tượng xã hội.....	33
II. NGÔN NGỮ VÀ TƯ DUY.....	33
IV. NGÔN NGỮ VÀ LỜI NÓI.....	35
V. CÁC CHỨC NĂNG CỦA NGÔN NGỮ.....	37
1. Chức năng giao tiếp.....	37
2. Chức năng phản ánh.....	37
3. Chức năng biểu cảm.....	38
VI. CÁC THUỘC TÍNH CỦA NGÔN NGỮ.....	39
1. Tính giao tiếp đối với tính thông tin.....	39
2. Thuộc tính duy nhất.....	39
3. Những thuộc tính khác.....	42
CHƯƠNG IV : NGÔN NGỮ LÀ HỆ THỐNG TÍN HIỆU.....	43
I. MỞ ĐẦU.....	43
II. TÍN HIỆU VÀ TÍN HIỆU NGÔN NGỮ.....	43

1. Tín hiệu.	43
2. Tín hiệu ngôn ngữ.	43
III. BẢN CHẤT CỦA TÍN HIỆU NGÔN NGỮ.	44
1. Ngôn ngữ phải có một dạng vật chất nhất định.	44
2. Tính vô đoán (arbitrary).	44
3. Tín hiệu ngôn ngữ là âm thanh.	44
IV. VỀ KHÁI NIỆM HỆ THỐNG VÀ CẤU TRÚC.	45
V. NGÔN NGỮ LÀ MỘT HỆ THỐNG CẤU TRÚC – CHỨC NĂNG.	45
1. Các đơn vị của hệ thống cấu trúc ngôn ngữ.	45
2. Các quan hệ của hệ thống cấu trúc ngôn ngữ.	46
VI. KẾT LUẬN.	47
PHẦN II : CƠ SỞ NGÔN NGỮ HỌC.	48
CHƯƠNG V : CƠ SỞ NGỮ ÂM HỌC VÀ ÂM VỊ HỌC.	48
I. NGỮ ÂM HỌC VÀ ÂM VỊ HỌC.	48
1. Giới thiệu.	48
2. Ngữ âm học thực nghiệm.	48
3. Ngữ âm học cấu âm.	48
4. Âm vị học.	49
5. Ngữ âm học thanh học.	49
6. Lịch sử.	50
II. CÁCH SẢN SINH ÂM TỔ LỜI NÓI.	50
1. Giới thiệu.	50
2. Cách sản sinh âm tố.	50
3. Các phụ âm và các nguyên âm.	50
4. Vị trí cấu âm và phương thức cấu âm.	51
5. Tiếng thanh.	51
6. Tính chất mũi (nasality).	52
7. Bảng chữ cái ngữ âm học quốc tế (The International Phonetic Alphabet).	52
III. MIÊU TẢ CÁC ÂM TỔ LỜI NÓI.	52
IV. MỘT SỐ HIỆN TƯỢNG NGỮ ÂM.	52
1. Các xu hướng phát âm (cấu âm).	52
2. Các quy luật ngữ âm.	53
3. Phiên âm và chuyển tự.	54
4. Các yếu tố điệu vị.	55
V. ÂM TIẾT VÀ CÁC LOẠI HÌNH ÂM TIẾT.	56
1. Khái niệm âm tiết.	56
2. Các loại hình âm tiết.	57
VI. ÂM VỊ VÀ CÁC KHÁI NIỆM CÓ LIÊN QUAN.	57
1. Khái niệm âm vị.	57
2. Âm vị, âm tố và biến thể của âm vị.	58
3. Thế đối lập âm vị học.	58
4. Phương pháp phân xuất âm vị và các biến thể của âm vị.	58
CHƯƠNG VI : CƠ SỞ TỪ VỰNG - NGỮ NGHĨA HỌC.	61
I. TỪ VỰNG VÀ TỪ VỰNG HỌC.	61
1. Từ vựng và các đơn vị từ vựng.	61
2. Từ vựng học.	63
II. NGỮ NGHĨA HỌC.	63

1. Giới thiệu.....	63
2. Những cách tiếp cận triết học.....	64
3. Những cách tiếp cận ngôn ngữ học.....	65
4. Ngữ nghĩa học đại cương.....	67
III. NGHĨA CỦA TỪ.....	67
1. Giới thiệu.....	67
2. Cấu trúc nghĩa của từ.....	68
3. Phân loại các từ về mặt nghĩa.....	69
4. Cơ cấu nghĩa của từ.....	69
5. Các quan hệ trong hệ thống từ vựng.....	71
CHƯƠNG VII : CƠ SỞ NGỮ PHÁP HỌC.....	75
I. NGỮ PHÁP VÀ NGỮ PHÁP HỌC.....	75
1. Khái niệm <i>ngữ pháp</i>	75
2. Đặc điểm của cơ cấu ngữ pháp.....	75
3. Ngữ pháp học.....	76
II. Ý NGHĨA VÀ PHƯƠNG THỨC NGỮ PHÁP.....	79
1. Ý nghĩa ngữ pháp.....	79
2. Phương thức ngữ pháp.....	80
III. PHẠM TRÙ NGỮ PHÁP.....	82
1. Giới thiệu.....	82
2. Phạm trù số.....	83
3. Phạm trù giống.....	83
4. Phạm trù cách.....	83
5. Phạm trù ngôi.....	84
6. Phạm trù thời.....	84
7. Phạm trù thể.....	84
8. Phạm trù thức.....	84
9. Phạm trù dạng.....	84
IV. HÌNH THÁI HỌC.....	85
1. Cấu trúc từ.....	85
2. Cách tạo từ.....	87
3. Hiện tượng biến tố.....	91
4. Từ loại.....	91
5. Hệ thống từ loại.....	94
V. CÚ PHÁP HỌC.....	97
1. Quan hệ cú pháp.....	97
2. Cụm từ.....	99
3. Câu.....	100
PHẦN III : CÁC NGÔN NGỮ CỦA THẾ GIỚI.....	103
CHƯƠNG VIII : VỀ VIỆC PHÂN LOẠI CÁC NGÔN NGỮ.....	103
I. GIỚI THIỆU.....	103
II. CÁC NGUYÊN TẮC PHÂN LOẠI NGÔN NGỮ.....	103
1. Khái niệm.....	103
2. Các nguyên tắc phân loại ngôn ngữ.....	104
III. CÁC PHƯƠNG PHÁP PHÂN LOẠI NGÔN NGỮ.....	104
1. Phương pháp so sánh - lịch sử.....	104
2. Phương pháp so sánh - loại hình.....	105

3. Phương pháp so sánh khu vực.....	105
4. Phương pháp so sánh-đối chiếu.....	105
IV. PHÂN LOẠI NGÔN NGỮ THEO NGUỒN GỐC.....	105
1. Việc xác lập những mối quan hệ họ hàng giữa các ngôn ngữ.....	106
2. Định luật Grimm và việc phục nguyên ngôn ngữ.....	107
V. PHÂN LOẠI NGÔN NGỮ VỀ MẶT LOẠI HÌNH.....	109
1. Loại hình và loại hình học ngôn ngữ.....	109
2. Phương pháp phân loại.....	110
3. Các ngôn ngữ theo loại hình học về hình thái.....	111
4. Các ngôn ngữ theo loại hình học về cú pháp.....	113
CHƯƠNG IX : CÁC NGÔN NGỮ CỦA THẾ GIỚI.....	117
I. CÁC NGÔN NGỮ ẤN – ÂU.....	117
1. Giới thiệu.....	117
2. Việc xác lập về ngữ hệ này.....	117
3. Sự tiến hóa.....	118
4. Nền văn hóa cổ xưa.....	118
II. CÁC NGÔN NGỮ CHÂU PHI.....	118
1. Giới thiệu.....	118
2. Phân loại các ngôn ngữ African.....	118
2. Hệ thống chữ viết châu Phi.....	122
3. Nghệ thuật của giao tiếp khẩu ngữ.....	123
III. CÁC NGÔN NGỮ AUSTRONESIAN.....	123
1. Giới thiệu.....	123
2. Phân loại.....	123
3. Các đặc trưng.....	123
IV. CÁC NGÔN NGỮ ASTRO-ASIATIC.....	124
V. CÁC NGÔN NGỮ HÁN-TẠNG.....	124
1. Giới thiệu.....	124
2. Tiểu ngữ hệ Trung Quốc.....	125
3. Tiểu ngữ hệ Tạng-Miến.....	125
4. Các nguồn gốc.....	125
5. Những đặc điểm ngôn ngữ.....	125
6. Các hệ thống chữ viết và văn học.....	125
7. Phân loại.....	126
VI. CÁC NGÔN NGỮ CAUCASIAN.....	126
1. Giới thiệu.....	126
2. Các nhóm ngôn ngữ.....	127
VII. CÁC NGÔN NGỮ FINO-UGRIC.....	127
1. Giới thiệu.....	127
2. Các nhóm ngôn ngữ.....	127
VIII. CÁC NGÔN NGỮ SEMITIC.....	128
1. Giới thiệu.....	128
2. Các nhóm ngôn ngữ.....	128
3. Các đặc điểm.....	128
4. Chữ viết.....	129
IX. CÁC NGÔN NGỮ ALTAIC.....	129
1. Giới thiệu.....	129

2. Các nhóm ngôn ngữ.....	129
X. CÁC NGÔN NGỮ URALIC.....	130
XI. CÁC NGÔN NGỮ ANH-ĐIÊNG MỸ.....	131
1. Giới thiệu.....	131
2. Các ngôn ngữ chính.....	131
3. Những đặc điểm của các ngôn ngữ Anh-điêng Mỹ.....	131
4. Các hệ thống chữ viết của ngôn ngữ Anh-điêng Mỹ.....	132
5. Những sự thêm các ngôn ngữ Anh-điêng vào tiếng Anh.....	132
6. Các ngôn ngữ Pidgins Anh-điêng Mỹ và những biệt ngữ thương mại.....	133
7. Ngôn ngữ kí hiệu của người Anh-điêng Mỹ.....	134
8. Phân loại.....	134
9. Các ngữ hệ ở Hoa Kỳ và Canada.....	135
LIỆU THAM KHẢO	136

PHẦN I : BẢN CHẤT XÃ HỘI CỦA NGÔN NGỮ

CHƯƠNG I : NGÔN NGỮ VÀ NGÔN NGỮ HỌC

I. NGÔN NGỮ.

1. Giới thiệu.

Ngôn ngữ (language) là phương tiện trọng yếu nhất được loài người sử dụng để giao tiếp với nhau. Ngôn ngữ về cơ bản là được nói, mặc dầu nó có thể được chuyển tải sang những phương tiện giao tiếp khác, chẳng hạn như viết. Nếu phương tiện của việc giao tiếp nói không có sẵn để dùng, như có thể là trường hợp giữa những người điếc, thì phương tiện trực quan chẳng hạn như *ngôn ngữ kí hiệu* (sign language) có thể được sử dụng. Một đặc trưng nổi bật của ngôn ngữ là ở chỗ mối quan hệ giữa một kí hiệu ngôn ngữ với ý nghĩa của nó luôn mang tính vô đoán (arbitrary), hoặc tính tùy tiện: không có lý do nào khác hơn là sự quy ước giữa những người nói tiếng Anh rằng một con chó cần phải được gọi là *dog*, và quả thực những ngôn ngữ khác nhau đều có những tên gọi khác nhau (ví dụ, tiếng Tây Ban Nha: *perro*, tiếng Nga: , tiếng Nhật: *inu*, tiếng Việt: *chó*). Ngôn ngữ có thể được sử dụng để bàn luận về một phạm vi rộng lớn của nhiều chủ đề, và đây là một đặc trưng phân biệt nó với giao tiếp động vật. Các điệu nhảy của những con ong mật, ví dụ, có thể được sử dụng chỉ để thông báo sự định vị của những nguồn thức ăn. Trong khi những khả năng học-ngôn ngữ về những sự bất chước nhiều đến ngạc nhiên - và điều này vẫn còn là sự bàn cãi vượt quá những giới hạn chính xác của những khả năng này, các nhà khoa học và các học giả nói chung đều đồng ý rằng những sự bất chước đó không tiến triển vượt quá những khả năng ngôn ngữ của một em bé hai tuổi.

2. Ngôn ngữ học.

Ngôn ngữ học (linguistics) là sự nghiên cứu khoa học về ngôn ngữ. Một số bộ môn của ngôn ngữ học sẽ được bàn luận dưới đây có liên quan đến những thành tố cơ bản của ngôn ngữ: *ngữ âm học* (phonetics) có liên quan đến những âm thanh của các ngôn ngữ, *âm vị học* (phonology) đánh dấu đến cách thức các âm thanh được sử dụng trong những ngôn ngữ riêng lẻ, *hình thái học* (morphology) đề cập đến cấu trúc của các từ, *cú pháp học* (syntax) có liên quan đến cấu trúc của những mệnh đề và câu, và *ngữ nghĩa học* (semantics) có quan hệ với sự nghiên cứu về ý nghĩa. Bộ môn cơ bản khác của ngôn ngữ học, *ngữ dụng học* (pragmatics), nghiên cứu sự tương tác giữa ngôn ngữ và các ngữ cảnh trong đó nó được sử dụng. *Ngôn ngữ học đồng đại* (synchronic linguistics) nghiên cứu hình thái của một ngôn ngữ tại một thời gian cố định trong lịch sử, đã qua hoặc hiện hữu. Trong khi đó, *ngôn ngữ học lịch đại* (diachronic linguistics), hoặc ngôn ngữ học lịch sử, lại khảo cứu cách thức và con đường mà một ngôn ngữ thay đổi qua thời gian. Một số lĩnh vực của ngôn ngữ học nghiên cứu các mối quan hệ giữa ngôn ngữ với sự kiện chủ đề của các bộ môn khoa học có liên quan, chẳng hạn như *ngôn ngữ học xã hội* (sociolinguistics) (xã hội học và ngôn ngữ) và *ngôn ngữ học tâm lý* (psycholinguistics) (tâm lý học và ngôn ngữ). Về mặt nguyên lý, *ngôn ngữ học ứng dụng* (applied linguistics) là việc ứng dụng bất kỳ về các phương pháp và những kết quả ngôn ngữ học để giải quyết những vấn đề có dính dáng đến ngôn ngữ, còn trong thực tế nó có xu hướng bị hạn định đối với sự chỉ dẫn ngôn ngữ-thứ hai.

3. Các thành phần của ngôn ngữ.

Ngôn ngữ con người được nói ra là phức thể của những âm thanh mà tự chúng không có nghĩa, nhưng những âm thanh này có thể được kết hợp với những âm thanh khác để tạo ra những thực thể có ý nghĩa. Chẳng hạn, trong tiếng Anh, các âm *p*, *e*, và *n* tự chúng chẳng có ý nghĩa gì, nhưng sự kết *pen* lại có ý nghĩa. Ngôn ngữ cũng được nêu đặc điểm bởi cú pháp phức tạp với những yếu tố, thường là các từ, được kết hợp vào trong những kết cấu phức tạp hơn, được gọi là những mệnh đề, và những kết cấu này chuyển sang đảm nhận một vai trò chính trong những cấu trúc của câu ra làm sao.

a. Các âm thanh của ngôn ngữ.

Vì hầu hết các ngôn ngữ được nói là chủ yếu, nên một phần quan trọng để hiểu biết đầy đủ về ngôn ngữ liên quan đến việc nghiên cứu về các âm thanh của ngôn ngữ.

Hầu hết âm thanh trong các ngôn ngữ của thế giới - và toàn bộ các âm thanh trong một số ngôn ngữ, chẳng hạn như tiếng Anh - được sản sinh ra bằng việc đẩy không khí từ phổi và thay đổi khoảng rộng phát âm giữa thanh quản và đôi môi. Ví dụ, âm *p* yêu cầu sự đóng chặt hoàn toàn của đôi môi, để không khí từ phổi đi ra bị bịt lại ở miệng, đưa lại việc hình thành nên đặc trưng âm nổ khi việc đóng kín môi được giải thoát. Đối với âm *s*, không khí từ phổi đi một cách liên tục qua miệng, nhưng lưỡi lại nâng lên gần như hoàn toàn tới nướu lợi (alveolar ridge) (khu vực của hàm trên chứa đựng các chân răng) để gây ra ma sát như là nó đóng khối từng phần luồng không khí đi qua đó. Các âm thanh cũng có thể được sản sinh bằng những phương thức khác hơn là việc tổng không khí từ phổi ra, và một số ngôn ngữ sử dụng những âm thanh này trong lời nói bình thường. Âm thanh được sử dụng bởi những người nói tiếng Anh để biểu thị sự quấy rầy, thường được đánh vần *tsk* hoặc *tut*, sử dụng không khí bị chặn lại trong khoảng không gian giữa mặt lưỡi, lưng lưỡi và vòm miệng. Những âm như vậy, được gọi là các âm mút (clicks), hoạt động như những âm tố bình thường trong các ngôn ngữ Khoisan của miền Tây Nam châu Phi và trong các ngôn ngữ Bantu của những người châu Phi lân cận.

Ngữ âm học là lĩnh vực sự nghiên cứu ngôn ngữ học quan tâm tới những thuộc tính vật lý của âm thanh, và nó có ba lĩnh vực: *ngữ âm học cấu âm* (articulatory) khảo sát thiết bị phát âm con người sản sinh ra các âm thanh như thế nào; *ngữ âm học thanh học* (acoustic phonetics) nghiên cứu các sóng âm được sản sinh ra bởi các thiết bị phát âm con người; và *ngữ âm học thính âm* (auditory phonetics) xem xét các âm thanh lời nói được lĩnh hội bởi tai con người ra làm sao. *Âm vị học* (phonology), trái lại, không quan tâm tới những thuộc tính vật lý của các âm thanh, mà đúng hơn là quan tâm chúng hoạt động trong một ngôn ngữ cụ thể như thế nào. Ví dụ sau đây minh họa sự khác nhau giữa ngữ âm học và âm vị học. Trong tiếng Anh, khi âm tố *k* (thông thường đánh vần là *c*) xuất hiện ở đầu một từ, như ở trong từ *cut*, nó được phát âm với *sự bật hơi* (aspiration). Tuy nhiên, khi âm này xuất hiện ở cuối một từ, như trong *tuck*, thì không có sự bật hơi. Về phương diện ngữ âm học, âm *k* bật hơi và âm *k* không bật hơi là những âm tố khác nhau, nhưng trong tiếng Anh, các âm tố khác nhau này chưa bao giờ khu biệt một từ này với một từ khác, và những người nói tiếng Anh thường không ý thức về sự khác nhau ngữ âm này cho đến khi nó được chỉ ra với họ. Như vậy, tiếng Anh không tạo ra sự khu biệt mang tính âm vị học giữa âm có bật hơi và âm không bật hơi *k*. Tiếng Hindi, trái lại, sử dụng sự khác nhau ngữ âm này để khu biệt những từ chẳng hạn như *kal* (thời gian), có một *k* không bật hơi, và *khal* (da), trong đó *kh* biểu hiện *k* bật hơi. Bởi vậy, trong tiếng Hindi, sự khu biệt giữa *k* bật hơi và không bật hơi là sự khu biệt cả về mặt ngữ âm học lẫn âm vị học.

b. Các đơn vị của ý nghĩa.

Trong khi nhiều người, do bị ảnh hưởng bởi việc viết, đều hướng suy nghĩ về các từ như là những đơn vị cơ bản của cấu trúc ngữ pháp, thì các nhà ngôn ngữ học lại tri nhận một đơn vị nhỏ hơn, *hình vị* (morpheme). Ví dụ, trong tiếng Anh, từ *cats* (những con mèo), gồm có hai yếu tố, hoặc hai hình vị, *cat*, nghĩa của nó có thể được nêu đặc trưng một cách thô thiển như là “động vật như mèo”, và *-s*, nghĩa của nó có thể được nêu đặc trưng một cách thô thiển như là “nhiều hơn một”. *Antimicrobial* (có tính kháng sinh), có nghĩa là “có năng lực về việc triệt phá các vi khuẩn” có thể được chia thành các hình vị *anti-* (chống lại), *microbe* (vi khuẩn), và *-ial*, một hậu tố cấu tạo từ này là một tính từ. Việc nghiên cứu về những đơn vị ngữ pháp nhỏ nhất này và các cách thức trong đó chúng kết hợp thành từ, được gọi là *hình thái học* (morphology).

c. Trật tự từ và cấu trúc câu.

Cú pháp học (syntax) là sự nghiên cứu về cách thức các từ kết hợp với nhau như thế nào để tạo thành câu. Trật tự của các từ trong các câu thay đổi từ ngôn ngữ này đến ngôn ngữ khác. Ví dụ, cú pháp ngôn ngữ Anh nói chung tuân theo trật tự chủ ngữ-động từ-bổ ngữ, như ở trong câu *The dog* (chủ ngữ) *bit* (động từ) *the man* (bổ ngữ). Câu *The dog the man bit* không phải là một kết cấu đúng trong tiếng Anh, và câu *The man bit*

the dog có một ý nghĩa rất khác biệt. Trái lại, tiếng Nhật có trật tự từ cơ bản là chủ ngữ-bổ ngữ-động từ, như ở trong *watakushi-wa hon-o kau*, mà dịch từng chữ là “Tôi quyển sách mua”. Tiếng Hixkaryana, được nói bởi khoảng 400 người tại một nhánh của dòng sông Amazon ở Brazil, có trật tự từ cơ bản là bổ ngữ-động từ-chủ ngữ. Câu *Toto yahosùye kamara*, dịch từng chữ là “Người đàn ông bị vỗ con báo đốm”, nghĩa thực tế là con báo đốm vỗ người đàn ông, chứ không phải là người đàn ông này vỗ con báo đốm.

Một đặc trưng chung của ngôn ngữ là ở chỗ các từ không được kết hợp một cách trực tiếp thành câu, mà là thành những đơn vị trung gian, được gọi là các *ngữ*, và rồi các ngữ được kết hợp thành câu. Câu *The shepherd found the lost sheep* (Shepherd tìm thấy con cừu bị mất) chứa đựng ít nhất ba ngữ: *the shepherd*, *found* và *the lost sheep*. Cấu trúc tôn ty này nhóm hợp các từ thành các ngữ, và các ngữ thành các câu, đảm nhiệm một vai trò quan trọng trong việc thiết lập những mối quan hệ với các câu này. Ví dụ, các ngữ *the shepherd* và *the lost sheep* xử sự như những đơn vị, do vậy khi câu này được tái phân bố thành dạng bị động, các đơn vị này vẫn không thay đổi: *The lost sheep was found by the shepherd*.

d. Nghĩa trong ngôn ngữ.

Trong khi các bộ môn về ngôn ngữ nghiên cứu hình thái của các yếu tố có liên quan đã được đề cập cơ bản trên đây, thì **ngữ nghĩa học** (semantics) là bộ môn của việc nghiên cứu có dính dáng đến nghĩa của những hình vị riêng biệt. Ngữ nghĩa học cũng còn liên quan đến việc nghiên cứu ý nghĩa của những kết cấu nối kết các hình vị để cấu tạo nên các ngữ và câu. Ví dụ, các câu *The dog bit the man* và *The man bit the dog* chứa đựng chính xác cùng các hình vị như nhau, nhưng chúng có những ý nghĩa khác nhau. Điều này là do các hình vị này tham gia vào những kết cấu khác nhau trong mỗi câu, được phản ánh trong những trật tự từ khác nhau của hai câu đó.

4. Việc thu đắc ngôn ngữ.

Việc *thu đắc ngôn ngữ* (language acquisition) còn được gọi là sự thu nhận ngôn ngữ, quá trình mà nhờ đó trẻ con và người lớn học một ngôn ngữ hoặc nhiều ngôn ngữ, là một lĩnh vực cơ bản của việc nghiên cứu ngôn ngữ.

a. Việc thu đắc ngôn ngữ thứ nhất.

Việc thu đắc ngôn ngữ thứ nhất (first-language acquisition) là một quá trình phức tạp mà các nhà ngôn ngữ học chỉ hiểu một cách cục bộ. Trẻ con có những đặc trưng bẩm sinh nhất định dẫn dắt chúng học ngôn ngữ. Những đặc trưng đó bao gồm cấu trúc của vùng phát âm, nó cho phép trẻ con tạo ra những âm thanh được sử dụng trong ngôn ngữ, và khả năng để hiểu một số lượng của các nguyên lý ngữ pháp chung, chẳng hạn như bản chất tôn ty về cú pháp. Tuy nhiên, những đặc trưng này không phải dẫn dắt trẻ con học chỉ một ngôn ngữ cụ thể. Trẻ con thu nhận bất cứ ngôn ngữ nào được nói xung quanh chúng, thậm chí nếu cha mẹ của chúng nói một ngôn ngữ riêng biệt. Một đặc tính thú vị của việc thu đắc ngôn ngữ sớm là ở chỗ trẻ con hình như tin cậy nhiều về ngữ nghĩa hơn là về cú pháp khi nói năng. Cái điểm mà tại đó chúng thay đổi đối với việc sử dụng cú pháp có vẻ là một điểm khấn yếu mà ở đó trẻ con có vẻ bất chước vượt trội hơn về khả năng ngôn ngữ.

b. Việc thu đắc ngôn ngữ thứ hai.

Mặc dù việc thu đắc ngôn ngữ thứ hai (second-language acquisition) quy chiếu về phương diện nghĩa câu chữ tới việc học một ngôn ngữ sau khi đã thu nhận được một ngôn ngữ thứ nhất, thì thuật ngữ này vẫn được sử dụng thường xuyên để quy chiếu đến việc thu đắc về một ngôn ngữ thứ hai sau khi một người đã đạt đến tuổi dậy thì. Trong khi trẻ con trải qua đôi chút khó khăn trong việc thu nhận nhiều hơn một ngôn ngữ, sau tuổi dậy thì người ta nói chung phải tiêu tốn công sức lớn hơn để học một ngôn ngữ thứ hai và họ thường đạt được những mức độ thấp hơn về *ngữ năng* (competence) trong ngôn ngữ đó. Nhiều người học các ngôn ngữ thứ hai thành công nhiều hơn khi họ trở nên đắm mình vào những nền văn hóa của những cộng đồng nói các ngôn ngữ đó. Nhiều người cũng học các ngôn ngữ thứ hai thành công nhiều trong những nền văn hóa mà

trong đó việc thu nhận một ngôn ngữ thứ hai được mong đợi, như ở đa số các nước châu Phi, hơn là họ học các ngôn ngữ thứ hai trong những nền văn hóa mà ở đó việc thành thạo ngôn ngữ thứ hai (second-language proficiency) được coi là khác thường, như trong đa số các nước nói tiếng Anh.

5. Hiện tượng song ngữ và đa ngữ.

Hiện tượng song ngữ (bilingualism) là khả năng làm chủ việc sử dụng hai ngôn ngữ, còn *hiện tượng đa ngữ* (multilingualism) là khả năng làm chủ việc sử dụng hai hoặc nhiều hơn hai ngôn ngữ. Mặc dù hiện tượng song ngữ tương đối hiếm giữa những người nói bản ngữ của tiếng Anh, nhưng trong nhiều bộ phận của thế giới, nó là cái tiêu chuẩn hơn là cái ngoại lệ. Ví dụ, hơn một nửa dân số của Papua New Guinea am hiểu theo phương diện chức năng về cả một ngôn ngữ bản xứ lẫn tiếng Tok Pisin. Nhiều người trong nhiều bộ phận của nước này đã làm chủ hai hoặc nhiều ngôn ngữ bản xứ. Hiện tượng song ngữ và hiện tượng đa ngữ liên quan đến những mức độ khác nhau về ngữ năng trong các ngôn ngữ có liên quan. Một người có thể điều khiển ngôn ngữ này tốt hơn ngôn ngữ khác, hoặc một người có thể làm chủ các ngôn ngữ khác nhau tốt hơn cho những mục đích khác nhau, ví dụ, trong việc sử dụng một ngôn ngữ này cho việc nói, còn ngôn ngữ khác cho việc viết.

6. Những sự đa dạng ngôn ngữ.

Các ngôn ngữ trải qua những sự thay đổi một cách triền miên, để lại kết quả trong sự phát triển của những biến thể khác nhau về các ngôn ngữ.

a. Các phương ngữ.

Một phương ngữ là một biến thể của một ngôn ngữ được nói bởi một tiểu nhóm người có thể nhận biết được. Về phương diện truyền thống, các nhà ngôn ngữ học đã áp dụng thuật ngữ *phương ngữ* (dialect) đối với những biến thể ngôn ngữ khu biệt về mặt địa lý, nhưng trong cách sử dụng hiện thời, thuật ngữ này có thể bao gồm cả những biến thể lời nói đặc trưng của những nhóm có thể xác định về phương diện xã hội khác. Việc xác định liệu hai biến thể lời nói là những phương ngữ của cùng một ngôn ngữ, hoặc liệu chúng đã có thay đổi đủ để coi là những ngôn ngữ khu biệt hay không, từng thường được chứng minh là một quyết định khó khăn và gây bàn cãi. Các nhà ngôn ngữ học thường viện dẫn việc có thể dễ hiểu lẫn nhau (mutual intelligibility) như là tiêu chuẩn chính trong việc thi hành quyết định này. Nếu hai biến thể lời nói không thể dễ hiểu lẫn nhau, thì những biến thể lời nói này là những ngôn ngữ khác nhau; nếu chúng có thể dễ hiểu lẫn nhau nhưng hơi khác về phương diện hệ thống với nhau, thì chúng là những phương ngữ của cùng một ngôn ngữ. Tuy nhiên, có những vấn đề với định nghĩa này, bởi vì nhiều mức độ của việc có thể dễ hiểu lẫn nhau tồn tại, và các nhà ngôn ngữ học cần phải quyết định tại cấp độ nào các biến thể lời nói có thể không còn được coi là việc có thể dễ hiểu lẫn nhau. Điều này khó xác lập trong thực tiễn. Tính có thể dễ hiểu lẫn nhau có một thành phần tâm lý học lớn: nếu một người nói của một biến thể lời nói này muốn hiểu một người nói của một biến thể lời nói khác, thì sự hiểu biết hình như có vẻ nhiều hơn nếu không phải là trường hợp này. Ngoài ra, các chuỗi của những sự đa dạng lời nói tồn tại trong đó những sự đa dạng lời nói kề bên có thể dễ hiểu lẫn nhau, nhưng những đa dạng lời nói xa hơn trong chuỗi này thì lại không. Hơn nữa, các nhân tố chính trị xã hội gần như tất yếu xen vào trong quá trình của việc phân biệt giữa các phương ngữ và ngôn ngữ. Những nhân tố như vậy, ví dụ, dẫn tới việc đặc trưng truyền hóa thống của tiếng Trung Quốc như một ngôn ngữ đơn lẻ với một số phương ngữ khó có thể dễ hiểu lẫn nhau.

Các phương ngữ phát triển chủ yếu như một kết quả của việc giao tiếp hạn chế giữa những bộ phận khác nhau của một cộng đồng chia phần một ngôn ngữ. Dưới những hoàn cảnh như vậy, những sự thay đổi xảy ra trong ngôn ngữ của một bộ phận cộng đồng không lan truyền sang nơi khác. Với tư cách là một kết quả, các biến thể lời nói này trở nên phân biệt nhiều hơn với biến thể lời nói khác. Nếu sự tiếp xúc tiếp tục bị hạn chế đối với một thời kỳ đủ dài, thì những thay đổi đầy đủ sẽ tích lũy để làm cho những biến thể lời nói khó có thể dễ hiểu lẫn nhau. Khi điều này xuất hiện, và đặc biệt nếu nó được phụ thêm bởi sự chia tách chính trị xã hội của một nhóm người nói từ cộng đồng lớn hơn, thì nó thường dẫn tới sự thừa nhận về những ngôn ngữ chia

tách. Những sự thay đổi khác nhau xảy ra trong tiếng La-tinh được nói ở những bộ phận khác nhau của Đế quốc La Mã, ví dụ, dần dần đưa đến sự nảy sinh các ngôn ngữ Romance hiện đại khác biệt, bao gồm tiếng Pháp, tiếng Tây Ban Nha, tiếng Bồ Đào Nha, tiếng Italia và tiếng Rumani.

Trong cách dùng bình thường, thuật ngữ *phương ngữ* có thể cũng biểu hiện một biến thể của một ngôn ngữ phân biệt với cái gì được xem xét hình thức chuẩn của ngôn ngữ đó. Tuy nhiên, các nhà ngôn ngữ học coi ngôn ngữ chuẩn đơn giản là một phương ngữ của một ngôn ngữ. Ví dụ, phương ngữ của tiếng Pháp được nói ở Paris trở thành ngôn ngữ chuẩn của tiếng Pháp không phải vì một số đặc thù ngôn ngữ của phương ngữ này, mà vì Paris là trung tâm văn hóa và chính trị của nước Pháp.

b. Những biến thể xã hội của ngôn ngữ.

Các *phương ngữ xã hội* (sociolects) là những phương ngữ được xác định bởi những nhân tố xã hội hơn là bởi địa lý. Các phương ngữ xã hội thường phát triển đồng thời với những sự phân chia xã hội trong một xã hội, chẳng hạn như những sự phân chia xã hội của tôn giáo và thành phần kinh tế xã hội. Ví dụ, ở New York City, xác suất mà người nào đó sẽ phát âm chữ cái *r* khi nó xuất hiện ở cuối của một âm tiết, như trong từ *fourth*, thay đổi với thành phần kinh tế xã hội. Cách phát âm về một *r* cuối nhìn chung được liên hệ với những thành viên các thành phần kinh tế xã hội cao hơn. Tình trạng như thế cũng đúng ở nước Anh về cách phát âm của *h*, như trong *hat*. Những thành viên của các nhóm xã hội nhất định thường chấp nhận một cách phát âm riêng như một cách tự phân biệt mình với những nhóm xã hội khác. Những người cư trú của Martha's Vineyard, Massachusetts, ví dụ, đã chấp nhận những cách phát âm nguyên âm riêng để tự phân biệt mình với những người đi nghỉ trên hòn đảo đó.

Slang, *argot*, và *jargon* là những thuật ngữ được chuyên môn hóa hơn cho các biến thể ngôn ngữ xã hội nhất định thường được xác định bởi những vốn từ vựng đặc biệt của chúng. *Slang* tham chiếu tới vốn từ vựng không chính thức, những cách đặt tên chết yểu, nó không thuộc về từ vựng chuẩn của ngôn ngữ. *Argot* quy chiếu một từ vựng phi chuẩn được sử dụng bởi những nhóm bí mật, những tổ chức tội phạm đặc biệt, thường có chủ định gây nên những sự giao tiếp khó hiểu cho những người ngoài. Một *jargon* gồm có từ vựng chuyên dụng của một giao dịch hoặc nghề nghiệp riêng, đặc biệt khi nó khó hiểu với những người ngoài, như với tiếng lóng hợp pháp.

Ngoài những biến thể ngôn ngữ được xác định trên cơ sở của các nhóm xã hội, có những biến thể ngôn ngữ được gọi là các *registers* (các ngôn ngữ mang tính quy ước) được xác định bởi tình huống xã hội. Ví dụ, trong một tình huống mang tính chính thức, một người có thể nói *You are requested to leave*, trong khi ở một tình huống không mang tính chính thức, cùng người đó có thể nói *Get out!*. Những sự khác nhau register có thể tác động đến cách phát âm, ngữ pháp và từ vựng.

c. Ngôn ngữ Pidgin và ngôn ngữ Creole.

Ngôn ngữ *pidgin* là một ngôn ngữ phụ trợ (auxiliary language) (một ngôn ngữ được sử dụng cho việc giao tiếp bởi những nhóm người có các tiếng bản ngữ khác nhau), nó phát triển khi mà những người nói các ngôn ngữ khác nhau buộc phải phát triển một phương tiện giao tiếp chung khi không có đủ thời gian để học một cách hoàn thiện các ngôn ngữ bản ngữ của nhau. Về mặt đặc trưng, một ngôn ngữ pidgin hầu hết vốn từ vựng của nó bắt nguồn từ một trong những ngôn ngữ bản ngữ này. Tuy nhiên, cấu trúc ngữ pháp của nó, sẽ hoặc là biến thể cao, phản ánh những cấu trúc ngữ pháp của mỗi ngôn ngữ bản ngữ của người nói, hoặc nó có thể theo thời gian trở nên được ổn định theo một phương thức rất khác với ngữ pháp của ngôn ngữ đã đóng góp hầu hết từ vựng cho nó. Về phương diện lịch sử, những xã hội đồn điền ở Caribbean và phía Nam Thái Bình Dương từng có nhiều ngôn ngữ pidgin. Tiếng Tok Pisin là ngôn ngữ pidgin chính của Papua New Guinea. Cả những nét tương đồng lẫn những nét dị biệt của nó với tiếng Anh có thể được nhìn thấy trong câu *Pik bilong dispela man i kam pinis*, nghĩa là *This man's pig has come*, hoặc, theo nghĩa câu chữ hơn là *Pig belong this-fellow man he come finish*.

Vì mỗi ngôn ngữ pidgin là một phụ ngôn ngữ phụ trợ, nên nó không có những người nói bản ngữ. Một ngôn ngữ *creole*, mặt khác, nảy sinh trong một tình huống tiếp xúc tương tự với tình huống sản sinh ra các ngôn ngữ pidgin và có lẽ đi qua một giai đoạn trong đó nó là một ngôn ngữ pidgin, nhưng một ngôn ngữ *creole* trở thành ngôn ngữ bản ngữ của cộng đồng của nó. Giống như với các ngôn ngữ pidgin, các ngôn ngữ *creole* thường lấy hầu hết từ vựng của chúng từ một ngôn ngữ đơn giản. Cũng giống như với các ngôn ngữ pidgin, cấu trúc ngữ pháp của một ngôn ngữ *creole* phản chiếu các cấu trúc của những ngôn ngữ, về mặt gốc gác trước đây được nói trong cộng đồng đó. Một đặc trưng của các ngôn ngữ *creole* là hình thái học đơn giản của chúng. Trong câu tiếng *creole* của người Jamaican *A fain Jan fain di kluoz*, nghĩa là *John tìm thấy quần áo*, từ vựng là gốc của tiếng Anh, trong khi cấu trúc ngữ pháp gấp đôi động từ để nhấn mạnh, phản ánh những mô hình ngôn ngữ Tây Phi. Vì các vốn từ của tiếng Pisin và tiếng *creole* Jamaican phần lớn từ gốc của tiếng Anh, nên chúng được gọi là ngôn ngữ *creole* có cơ sở tiếng Anh (English-based creple).

7. Các ngôn ngữ của thế giới.

Những đánh giá về số lượng của các ngôn ngữ được nói trên thế giới ngày nay thay đổi phụ thuộc vào ranh giới phân chia (dividing line) giữa ngôn ngữ và các phương ngữ được vạch ra ở đâu. Ví dụ, các nhà ngôn ngữ học bất đồng ý kiến liệu tiếng Trung Quốc có cần phải được coi là một ngôn ngữ riêng biệt vì truyền thống văn học và nền văn hóa được dùng chung của những người nói nó, hoặc liệu nó có cần phải được coi là một vài ngôn ngữ khác nhau vì tính khó có thể dễ hiểu lẫn nhau, ví dụ, về tiếng Quan thoại (Mandarin) được nói ở Bắc Kinh với tiếng Quảng Đông được nói ở Hồng Kông hay không. Nếu tính có thể dễ hiểu lẫn nhau là tiêu chuẩn cơ bản, thì những đánh giá hiện thời chỉ định rằng có khoảng 6000 ngôn ngữ được nói trên thế giới ngày nay. Tuy nhiên, nhiều ngôn ngữ với một số lượng người nói nhỏ hơn đang có nguy cơ bị thay thế bởi những ngôn ngữ với số lượng người nói lớn hơn. Trên thực tế, một số học giả tin rằng có lẽ 90% của những ngôn ngữ được nói vào những năm 1990 sẽ tắt hoặc bị chết với sự tàn lụi vào khoảng cuối thế kỷ 21. 12 ngôn ngữ được nói rộng rãi nhất, với số lượng xấp xỉ về những người nói bản ngữ, là như sau: Tiếng Trung Quốc Quan thoại, 836 triệu; Tiếng Ấn Độ, 333 triệu; Tiếng Tây Ban Nha, 332 triệu; tiếng Anh, 322 triệu; tiếng Bengali, 189 triệu; tiếng Ả Rập, 186 triệu; tiếng Nga, 170 triệu; tiếng Bồ Đào Nha, 170 triệu; tiếng Nhật, 125 triệu; tiếng Đức, 98 triệu; tiếng Pháp, 72 triệu; tiếng Malaysia, 50 triệu. Nếu những người nói ngôn ngữ-thứ hai được gộp vào đây, thì tiếng Anh là ngôn ngữ thứ hai được nói rộng rãi nhất, với 418 triệu người nói.

a. Phân loại các ngôn ngữ.

Các nhà ngôn ngữ học phân loại các ngôn ngữ trong khi sử dụng hai hệ thống phân loại chính: phân loại loại hình và phân loại cội nguồn. Hệ thống phân loại loại hình tổ chức các ngôn ngữ theo những nét giống nhau lẫn những nét khác nhau về các cấu trúc của chúng. Các ngôn ngữ chia sẻ cùng cấu trúc sẽ thuộc về cùng loại hình, trong khi những ngôn ngữ với những cấu trúc khác nhau sẽ thuộc về những loại hình khác nhau. Ví dụ, dù những nét khác nhau lớn giữa hai ngôn ngữ theo những phương diện khác, nhưng tiếng Trung Quốc Quan thoại và tiếng Anh thuộc về cùng loại hình, được nhóm hợp lại theo loại hình học trật tự từ (word-order typology). Cả hai ngôn ngữ có một trật tự từ cơ bản về chủ ngữ-động từ-bổ ngữ.

Sự phân loại cội nguồn về các ngôn ngữ phân chia chúng thành những *ngữ hệ* (language family) trên cơ sở về sự phát triển lịch sử của chúng: một nhóm các ngôn ngữ cùng nguồn gốc (descend) về phương diện lịch sử từ cùng một tổ tiên chung cấu thành nên một ngữ hệ. Ví dụ, các ngôn ngữ Romance cấu thành một ngữ hệ bởi vì chúng hoàn toàn có nguồn gốc từ ngôn ngữ La-tinh. Tiếng La-tinh, đến lượt mình, lại thuộc về ngữ hệ lớn hơn, ngữ hệ Ấn-Âu, mà ngôn ngữ tổ tiên của nó được gọi là ngôn ngữ Tiên Ấn-Âu (Proto-Indo-European). Một số việc nhóm hợp cội nguồn được chấp nhận một cách phổ quát. Tuy nhiên, vì những tài liệu làm chứng đối với hình thái của đa số các ngôn ngữ tổ tiên, kể cả ngôn ngữ Tiên Ấn-Âu, đã không sống sót, nên gây nhiều tranh cãi xung quanh những cách nhóm hợp cội nguồn phạm vi rộng hơn. Một sự khảo sát mang tính dè dặt về các ngữ hệ của thế giới là như ở dưới đây.

b. Ngữ hệ Ấn-Âu.

Các ngôn ngữ Ấn-Âu là những ngôn ngữ được nói rộng rãi nhất ở châu Âu, và chúng cũng trải dài sang Nam và Tây châu Á. Ngữ hệ này bao gồm một số lượng của các tiểu ngữ hệ hoặc các nhánh (những nhóm ngôn ngữ có cội nguồn từ một tổ tiên chung, đến lượt mình nó là một thành viên của một nhóm ngôn ngữ lớn hơn (có nguồn gốc từ một tổ tiên chung khác). Hầu hết những người ở vùng Tây Bắc châu Âu nói các ngôn ngữ Romance, bao gồm tiếng Anh, tiếng Đức, và tiếng Hà Lan cũng như các ngôn ngữ Scandinavian, chẳng hạn như tiếng Đan Mạch, tiếng Na Uy và tiếng Thụy Điển. Các ngôn ngữ Celtic, chẳng hạn như tiếng Welsh và tiếng Gaelic, một lần đã từng bao trùm một phần rộng của châu Âu, nhưng bây giờ đã bị hạn chế tới những đường viền phía Tây của nó. Các ngôn ngữ Romance tất cả đều có nguồn gốc từ tiếng La-tinh, là những ngôn ngữ duy nhất còn sống sót của một ngữ hệ rộng lớn hơn, mà nó bao gồm, ngoài tiếng La-tinh, một số ngôn ngữ bây giờ đã tắt của tiếng Italy. Các ngôn ngữ của nhánh Baltic và Slavic (slavonic) có quan hệ một cách gần gũi. Chỉ hai trong các ngôn ngữ Baltic sống sót: tiếng Lithuanian và tiếng Latvian. Các ngôn ngữ Slavic, bao trùm nhiều phần Trung và Đông Âu, gồm có tiếng Nga, tiếng Ukrainian, tiếng Ba Lan, tiếng Tiệp, tiếng Serbo – Croatian và tiếng Bulgarian. Trên bán đảo Balkan, hai nhánh của ngữ hệ Ấn-Âu tồn tại mà mỗi nhánh gồm có một ngôn ngữ đơn lẻ - ấy là tiếng Hy Lạp và tiếng Albanian. Xa hơn về phía Đông, ở Caucasia, ngôn ngữ Armenian cấu thành nên một nhánh ngôn ngữ đơn lẻ khác của ngữ hệ Ấn-Âu.

Nhánh chính còn sót lại khác của ngữ hệ Ấn-Âu là nhánh Ấn Độ-Iran. Nó có hai tiểu nhánh, tiểu nhánh Iranian và tiểu nhánh Indo-Aryan (Indic). Các ngôn ngữ của tiểu nhánh Iran được nói chủ yếu ở Tây Nam châu Á và bao gồm tiếng Ba Tư, tiếng Pashto (được nói ở Afghanistan) và tiếng Kurdish. Các ngôn ngữ Indo-Aryan được nói ở bộ phận phía Bắc của Nam Á (Pakistan, Bắc Ấn Độ, Nepal và Bangladesh) và cũng ở hầu hết Sri Lanka. Tiểu nhánh này bao gồm tiếng Ấn Độ-tiếng Urdu, tiếng Bengali, tiếng Nepali, và tiếng Sinhalese (ngôn ngữ được nói bởi phần lớn người dân ở Sri Lanka). Các tài liệu lịch sử đã xác nhận về những nhánh khác của ngữ hệ Ấn-Âu, chúng bây giờ đã tắt, chẳng hạn như các ngôn ngữ Anatolian, chúng đã từng một lần được nói ở cái mà hiện nay được gọi là Thổ Nhĩ Kỳ và bao gồm ngôn ngữ Hittite cổ đại.

c. Các ngữ hệ Ấn-Âu khác.

Các ngôn ngữ Uralic cấu thành nên một ngữ hệ Ấn-Âu cơ bản khác. Chúng được nói phần lớn ở phần Đông Bắc của lục địa này, tràn phủ qua Tây Bắc châu Á; một ngôn ngữ, tiếng Hunggari, được nói ở Trung Âu. Đa số các ngôn ngữ Uralic thuộc về nhánh Finno-Ugric của ngữ hệ này. Nhánh này bao gồm (ngoài tiếng Hunggari) tiếng Phần Lan, tiếng Estonian và tiếng Saami. Châu Âu cũng có một ngôn ngữ biệt lập (một ngôn ngữ chưa biết được là có liên quan tới bất cứ ngôn ngữ nào): tiếng Basque, được nói ở Pyrenees. Dãy núi Caucasus nằm tại ranh giới giữa Đông Nam châu Âu và châu Á. Từ các thời kỳ cổ xưa, vùng này đã từng chứa đựng một số lượng lớn các ngôn ngữ, bao gồm hai nhóm ngôn ngữ, về mặt xác định không có liên quan đến các ngữ hệ bất kỳ khác. Tiếng Caucasus Nam, hoặc tiếng Kartvelian, những ngôn ngữ được nói ở Georgia và kể cả ngôn ngữ Georgian. Các ngôn ngữ Caucasus Bắc chia thành tiểu nhóm Caucasus Tây Bắc, tiểu nhóm Caucasus Trung Bắc và tiểu Caucasus Đông Bắc. Quan hệ cội nguồn của tiểu nhóm Caucasus Tây Bắc với những tiểu nhóm chưa được đồng ý về phương diện phổ quát. Các ngôn ngữ Caucasus Tây Bắc bao gồm tiếng Abkhaz, các ngôn ngữ Caucasus Trung Bắc gồm tiếng Chechen, và các ngôn ngữ Caucasus Đông Bắc bao gồm ngôn ngữ Avar.

d. Các ngữ hệ châu Á và Thái Bình Dương.

Ngoài nhánh Indo-Aryan của ngữ hệ Ấn-Âu, Nam Á chứa đựng hai ngữ hệ lớn khác. Ngữ hệ Dravidian nổi trội ở Nam Ấn Độ và bao gồm tiếng Tamil lẫn tiếng Telugu. Các ngôn ngữ Munda đại diện ngữ hệ Austro-Asiatic ở Ấn Độ và chứa đựng nhiều ngôn ngữ, mỗi ngôn ngữ có lượng người nói tương đối nhỏ. Ngữ hệ Austro-Asiatic cũng lan sang Đông Nam Á, nơi nó bao gồm tiếng Khmer (tiếng Chăm-pu-chia) và tiếng Việt. Nam Á chứa đựng ít nhất một ngôn ngữ biệt lập, tiếng Burushaski, được nói ở cách xa phía Bắc của Pakistan.

Một số nhà ngôn ngữ học tin rằng nhiều ngôn ngữ của miền Trung, miền Bắc và miền Đông Á cấu tạo nên một ngữ hệ Altaic đơn lẻ, cho dù một số nhà ngôn ngữ học khác coi tiếng Turkic, tiếng Tungusic và tiếng Mongolic là những ngữ hệ tách biệt, không có quan hệ. Các ngôn ngữ Turkic bao gồm tiếng Thổ Nhĩ Kỳ và một số lượng ngôn ngữ của cựu Liên bang Cộng hòa Xã hội chủ nghĩa Xô Viết (Liên Xô), chẳng hạn như tiếng Uzbek và tiếng Tatar. Các ngôn ngữ Tungusic được nói chủ yếu bởi những nhóm nhỏ dân cư ở Siberia và Đông Bắc Trung Quốc. Ngữ hệ này bao gồm ngôn ngữ Manchu đã tắt mới đây. Ngôn ngữ chính của ngữ hệ Mongolic là tiếng Mông Cổ. Một số nhà ngôn ngữ học cũng gán tiếng Nhật và tiếng Triều Tiên cho ngữ hệ Altaic, mặc dầu nhiều nhà ngôn ngữ học khác lưu tâm tới những ngôn ngữ này như là những ngôn ngữ biệt lập. Ở Bắc Á có một số ngôn ngữ xuất hiện hoặc hình thành nên những ngữ hệ nhỏ, độc lập hoặc là những ngôn ngữ biệt lập, chẳng hạn như ngữ hệ Chukotko – Kamchatkan của những bán đảo Chukot và Kamchatka ở vùng viễn đông của Nga. Các ngôn ngữ này thường được quy chiếu mang tính tập hợp như là nhóm Paleo-Siberian (hay còn gọi là Paleo-Asiatic), nhưng đây là một sự nhóm họp địa lý, chứ không phải là một sự nhóm họp cội nguồn.

Ngữ hệ Hán-Tạng bao phủ không chỉ hầu hết Trung Quốc, mà còn nhiều phần dãy Himalayas và nhiều phần của Đông Nam Á. Các ngôn ngữ chính của ngữ hệ này là tiếng Trung Quốc, tiếng Tạng và tiếng Miến Điện. Các ngôn ngữ Tày (Tai) cấu thành nên một ngữ hệ quan trọng khác của Đông Nam Á. Chúng được nói ở Thailand, Lào và Nam Trung Quốc và kể cả tiếng Thái. Các ngôn ngữ Mèo – Dao, hoặc Hmong-Mien, được nói trong những vùng cô lập của Nam Trung Quốc và Bắc Đông Nam Á. Các ngôn ngữ Austronesian, trước đây được gọi là các ngôn ngữ Malayo-Polynesian, bao trùm Bán đảo Mã Lai và hầu hết các quần đảo Đông Nam châu Á và được nói ở vùng viễn tây Madagascar và khắp suốt những hòn đảo Thái Bình Dương cũng như vùng viễn đông Easter Island. Các ngôn ngữ Austronesian bao gồm tiếng Malaysia (được gọi là Bahasa Malaysia ở Malaysia, và Bahasa Indonesia ở Indonesia), tiếng Java, tiếng Hawaiian và tiếng Maori (ngôn ngữ của những người nguyên thủy New Zealand).

Mặc dầu những người cư trú của một số vùng bờ biển và các quần đảo New Guinea nói các ngôn ngữ Austronesian, nhưng hầu hết cư dân của hòn đảo chính, cũng như một số cư dân của những hòn đảo kế cận nói các ngôn ngữ không có quan hệ với các tiếng Austronesian. Các nhà ngôn ngữ học đã quy chiếu về phương diện tập hợp những ngôn ngữ này như là những ngôn ngữ Papuan, mặc dầu đây là một thuật ngữ địa lý bao trùm khoảng 60 ngữ hệ khác nhau. Các ngôn ngữ của những thổ dân Aborigines Úc cấu thành nên nhóm không có quan hệ khác, và quả là có thể tranh luận liệu tất cả các ngôn ngữ Úc cấu thành nên một ngữ hệ đơn lẻ hay không.

e. Các ngữ hệ châu Phi.

Các ngôn ngữ của châu Phi có thể thuộc về bốn ngữ hệ: ngữ hệ Afro-Asiatic, ngữ hệ Nilo-Saharan, ngữ hệ Niger-Congo, và ngữ hệ Khoisan, cho dù tính đồng nhất cội nguồn của ngữ hệ Nilo-Saharan và ngữ hệ Khoisan vẫn còn đang bàn cãi. Các ngôn ngữ Afro-Asiatic chiếm giữ hầu hết Bắc Phi và cũng nhiều phần rộng lớn của Tây Nam Á. Ngữ hệ này gồm có vài nhánh. Nhánh Semitic bao gồm tiếng Ả Rập, tiếng Do Thái và nhiều ngôn ngữ của Ethiopia lẫn Eritrea, kể cả tiếng Amharic, ngôn ngữ nổi trội của Ethiopia. Nhánh Chadic, được nói chủ yếu ở Bắc Nigeria và những vùng kề cận, bao gồm tiếng Hausa, một trong hai ngôn ngữ được nói rộng rãi nhất của miền dưới Saharan châu Phi (hiện thân tiếng Swahili khác). Những tiểu ngữ hệ khác của ngữ hệ Afro-Asiatic là tiểu ngữ hệ Berber, tiểu ngữ hệ Cushitic, và ngôn ngữ nhánh đơn lẻ tiếng Ai Cập, chứa đựng ngôn ngữ hiện nay đã tắt của những người Ai Cập cổ đại.

Ngữ hệ Niger-Congo bao trùm hầu hết miền dưới Saharan châu Phi và bao gồm những ngôn ngữ Tây Phi được nói rộng rãi chẳng hạn như tiếng Yoruba và tiếng Fulfulde, cũng như những ngôn ngữ Bantu của Nam và Đông Phi, bao gồm tiếng Swahili và tiếng Zulu. Các ngôn ngữ Nilo-Saharan được nói chủ yếu ở Đông Phi, trong một vùng bị bao phủ giữa các ngôn ngữ Afro-Asiatic và Niger-Congo. Ngôn ngữ Nilo-Saharan được biết rõ nhất là tiếng Masai, được nói bởi những người Masai ở Kenya và Tanzania. Các ngôn

ngữ Khoisan được nói ở góc Tây Nam của châu Phi và bao gồm ngôn ngữ Nama (trước đây được gọi là tiếng Hottentot).

f. Các ngữ hệ của châu Mỹ.

Một số nhà ngôn ngữ học nhóm họp tất cả các ngôn ngữ bản xứ của châu Mỹ vào trong chỉ ba ngữ hệ, trong khi đa số tách riêng chúng thành một số lượng lớn các ngữ hệ và các ngôn ngữ biệt lập. Các ngữ hệ được xác lập một cách rõ ràng nhất bao gồm ngữ hệ Eskimo - Aleut. Ngữ hệ này trải rộng từ miền Đông Siberia đến quần đảo Aleutian, và ngang qua Alaska và Bắc Canada tới Greenland, nơi mà một biến thể của ngôn ngữ Inuit, tiếng Greenlandic, là một ngôn ngữ hành chính. Các ngôn ngữ Na-Dené, nhánh chính của nó gồm có các ngôn ngữ Athapaskan, chiếm giữ phần lớn miền Tây Bắc Bắc Mỹ. Tuy nhiên, các ngôn ngữ Athapaskan cũng bao gồm một nhóm ngôn ngữ ở Tây Nam nước Mỹ, một trong số chúng là tiếng Navajo. Các ngôn ngữ của ngữ hệ Algonquian và ngữ hệ Iroquoian cấu thành nên những ngôn ngữ bản xứ chính của Đông Bắc Bắc Mỹ, trong khi ngữ hệ Siouan là một trong những ngữ hệ chính của miền Trung Bắc Mỹ.

Ngữ hệ Uto-Aztecan mở rộng từ Tây Nam Hoa Kỳ sang miền Trung châu Mỹ và bao gồm tiếng Nahuatl, ngôn ngữ của nền văn minh Aztec và các con cháu hiện đại của nó. Các ngôn ngữ Mayan được nói chủ yếu ở Mexico và miền Nam Guatemala. Các ngữ hệ chính của Nam Mỹ bao gồm các ngữ hệ Carib và Arawak ở phía Bắc, Macro-Gê và Tupian Gê ở phía Đông. Tiếng Guaraní, được thừa nhận như một ngôn ngữ quốc gia ở Paraguay bên cạnh ngôn ngữ hành chính, tiếng Tây Ban Nha, là một thành viên quan trọng của ngữ hệ Tupian. Ở vùng dãy núi Andes, các ngôn ngữ bản xứ nổi trội là tiếng Quechua và tiếng Aymara; quan hệ cội nguồn của những ngôn ngữ này với nhau và với các ngôn ngữ khác vẫn còn gây tranh cãi.

g. Việc phục nguyên các ngôn ngữ.

Phục nguyên ngôn ngữ là sự khôi phục về những giai đoạn của một ngôn ngữ đã từng tồn tại trước những cái được tìm thấy bên trong các tài liệu viết. Trong khi sử dụng một số lượng của những ngôn ngữ có quan hệ về phương diện cội nguồn, các nhà ngôn ngữ học cố gắng phục nguyên ít ra là những phương diện cơ bản của một ngôn ngữ tổ tiên chung của các ngôn ngữ, được gọi là *ngôn ngữ tiền thân* (protolanguage). Các nhà ngôn ngữ học lập luận rằng những thuộc tính đó là như nhau giữa những ngôn ngữ con cháu của ngôn ngữ tiền thân protolanguage, hoặc những thuộc tính đó là khác nhau nhưng có thể để lại vết tích về một gốc gác chung, thì có thể được xem là những thuộc tính của ngôn ngữ tổ tiên. Khoa học ngôn ngữ học thế kỷ thứ 19 đã thiết lập một quá trình quan trọng, có ý nghĩa trong việc phục nguyên ngôn ngữ Tiền Ấn-Âu (Proto-Indo-European language). Trong khi nhiều chi tiết của việc phục nguyên này còn gây bàn cãi, thì nhìn chung các nhà ngôn ngữ học đã thu được một nhận thức tốt về âm vị học, hình thái học và từ vựng của ngôn ngữ Tiền Ấn-Âu. Tuy nhiên, vì phạm vi của sự biến đổi cú pháp giữa các ngôn ngữ con cháu của tiếng Tiền Ấn-Âu, các nhà ngôn ngữ học đã nhận thấy việc phục nguyên cú pháp là việc phục nguyên có nhiều nghi vấn hơn cả.

i. Các ngôn ngữ Pidgin và Creole.

Các ngôn ngữ pidgin và creole riêng lẻ nêu ra một vấn đề đặc biệt cho việc phân loại cội nguồn bởi vì từ vựng và ngữ pháp của mỗi một ngôn ngữ này bắt nguồn từ những nguồn khác nhau. Về mặt hệ quả, nhiều nhà ngôn ngữ học không cố gắng phân loại chúng về phương diện cội nguồn. Các ngôn ngữ pidgin và creole được tìm thấy trong nhiều phần của thế giới, nhưng có những sự tập trung đặc biệt ở Caribbean, Tây Phi, và nhiều quần đảo của Ấn Độ Dương lẫn Nam Thái Bình Dương. Các ngôn ngữ creole được dựa trên tiếng Anh (English-based creoles) chẳng hạn như ngôn ngữ Creole Jamaican và Creole Guyanese, và các ngôn ngữ creole được dựa trên tiếng Pháp (French-based creoles) chẳng hạn như ngôn ngữ Creole Haitian, có thể được tìm thấy ở Caribbean. Các ngôn ngữ creole được dựa trên tiếng Anh lan rộng ở Tây Phi. Khoảng 10% dân số của Sierra Leone nói tiếng Krio như một ngôn ngữ bản địa, và thêm 85% nói nó như một ngôn ngữ thứ hai. Các ngôn ngữ creole của những quần đảo Ấn Độ Dương, chẳng hạn như Mauritius lại dựa trên tiếng Pháp. Một ngôn ngữ creole dựa trên tiếng Anh, tiếng Tok Pisin, được nói bởi hơn 2 triệu người ở Papua

New Guinea, làm cho nó trở thành ngôn ngữ phụ được nói rộng rãi nhất của nước này. Các cư dân của Solomon Islands và Vanuatu nói những biến thể tương tự của tiếng Tok Pisin, được gọi là tiếng Pijin và tiếng Bislama, tương ứng từng cái một.

k. Các ngôn ngữ quốc tế.

Các ngôn ngữ quốc tế bao gồm cả những ngôn ngữ đang tồn tại đã trở thành phương tiện giao tiếp quốc tế lẫn những ngôn ngữ nhân tạo được xây dựng để phục vụ mục đích này. Ngôn ngữ quốc tế nhân tạo nổi tiếng và phổ cập nhất là *Quốc tế ngữ* (Esperanto); tuy nhiên, các ngôn ngữ quốc tế lan rộng nhất không phải là ngôn ngữ nhân tạo. Ở châu Âu thời Trung cổ, tiếng La-tinh là ngôn ngữ quốc tế thiết yếu. Ngày nay, tiếng Anh được sử dụng ở nhiều nước như một ngôn ngữ hành chính hoặc như một phương tiện chính của giao tiếp quốc tế hơn mọi ngôn ngữ khác. Tiếng Pháp là ngôn ngữ được sử dụng rộng rãi thứ hai, đồng thời là ngôn ngữ hành chính của phần lớn các nước châu Phi. Các ngôn ngữ khác được sử dụng hạn chế mang tính khu vực hơn, chẳng hạn như tiếng Tây Ban Nha ở Tây Ban Nha và Mỹ La-tinh, tiếng Ả Rập ở Trung Đông, và tiếng Nga ở những nước cộng hòa của Liên Xô cũ.

8. Ngôn ngữ phi lời nói.

Ngôn ngữ, mặc dù nói là chủ yếu, có thể cũng được biểu hiện trong phương tiện truyền thông (media) khác, chẳng hạn như viết. Dưới những chu cảnh nhất định, ngôn ngữ nói có thể bị hắt cẳng bởi phương tiện truyền thông khác, như trong ngôn ngữ kí hiệu (sign language) giữa những người điếc. Ngôn ngữ viết có thể được nhìn nhận theo một ý nghĩa như là sự ghi chép vật lý lâu dài hơn (permanent) của ngôn ngữ nói. Tuy nhiên, ngôn ngữ viết và ngôn ngữ nói hình có khuynh hướng đi rẽ ra khỏi nhau, một phần là vì sự khác nhau về môi trường. Trong ngôn ngữ nói, cấu trúc của một thông điệp không thể quá phức tạp bởi vì sự mạo hiểm mà người nghe sẽ hiểu lầm thông báo. Tuy nhiên, vì giao tiếp là mặt đối mặt (face-to face) nên người nói có cơ hội để tiếp nhận sự phản hồi của người nghe và làm rõ ràng điều mà người nghe nào hiểu. Các cấu trúc câu trong giao tiếp viết có thể phức tạp hơn vì người đọc có thể quay trở lại một bộ phận trước của văn bản để làm rõ về việc hiểu của mình. Tuy nhiên, người viết thường không có cơ hội để tiếp nhận sự phản hồi của người đọc và để tái lập (rework) văn bản, do vậy các văn bản phải được viết với sự rõ ràng lớn hơn. Một ví dụ về sự khác nhau này giữa ngôn ngữ viết và ngôn ngữ nói được tìm thấy trong những ngôn ngữ có những biến thể chữ viết chỉ mới được phát triển gần đây. Trong những biến thể viết này có sự tăng nhanh việc sử dụng của các từ chẳng hạn như *vì* và *tuy nhiên* để làm rõ những mối liên kết giữa các câu - những mối liên kết mà bình thường bị để ẩn trong ngôn ngữ nói.

Các ngôn ngữ kí hiệu, khác với những phiên bản được ký hiệu của các ngôn ngữ nói, là những ngôn ngữ chung (native languages) của đa số thành viên trong những cộng đồng người điếc. Các nhà ngôn ngữ học chỉ gần đây mới bắt đầu đánh giá những mức độ về tính phức tạp và tính diễn cảm được tìm thấy trong các ngôn ngữ kí hiệu. Cụ thể, cũng giống như trong các ngôn ngữ miệng, các ngôn ngữ kí hiệu nói chung là có tính vô đoán trong việc sử dụng của chúng về các kí hiệu: nói chung, chẳng có lý gì khác ngoài sự quy ước (convention) cho một kí hiệu nhất định có một ý nghĩa cụ thể. Các ngôn ngữ kí hiệu cũng phô bày việc lập mô hình kép, trong đó một số nhỏ của các thành tố kết hợp để sản sinh ra phạm vi tổng thể về các kí hiệu, tương tự như cách thức mà ở đó các kí tự kết hợp để tạo ra các từ trong tiếng Anh. Ngoài ra, các ngôn ngữ kí hiệu sử dụng cú pháp phức tạp và có thể bàn luận về cùng phạm vi rộng của những chủ đề có thể có trong những ngôn ngữ nói.

Ngôn ngữ cơ thể (body language) quy chiếu việc chuyển tải về những thông điệp thông qua những sự vận động cơ thể khác hơn là những sự vận động cấu thành nên một phần của của ngôn ngữ kí hiệu hoặc ngôn ngữ nói. Một vài cử chỉ có thể có những có nghĩa hoàn toàn đặc biệt, chẳng hạn như cử chỉ để nói tạm biệt hoặc để yêu cầu tiếp cận một người nào đó. Những cử chỉ khác nói chung đi kèm lời nói, chẳng hạn như những cử chỉ được sử dụng để nhấn mạnh một điểm cụ thể. Mặc dù có những nét tương đồng xuyên văn hóa trong ngôn ngữ cơ thể, nhưng những nét dị biệt mang tính chủ quan cũng tồn tại cả trong phạm vi mà ngôn

ngữ cơ thể được sử dụng lẫn cả trong những cách giải thích được cho sẵn đối với những trường hợp cụ thể về ngôn ngữ cơ thể. Ví dụ, các cử chỉ của đầu cho “có” và “không” được sử dụng trong những người Balkans có vẻ bị đảo ngược với những người châu Âu khác. Đồng thời, khoảng cách vật lý được chiếm giữ giữa những người tham gia trong một cuộc nói chuyện thay đổi từ nền văn hóa này đến nền văn hóa khác: một khoảng cách được coi là bình thường trong một văn hóa này có thể đánh bật một người nào đó từ nền văn hóa khác với tư cách là một giới hạn công kích.

Trong những chu cảnh nhất định, phương tiện truyền thông (media) khác có thể được sử dụng để chuyên chở những thông điệp ngôn ngữ, đặc biệt khi mà phương tiện truyền thông bình thường không có sẵn. Ví dụ, mã Morse mã hóa một cách trực tiếp một thông điệp viết, chữ theo chữ, do vậy mà nó có thể được truyền đi bởi một môi trường cho phép chỉ có hai giá trị – về mặt truyền thống – các tín hiệu ngắn và dài hoặc các dấu chấm nhỏ (dots) và các dấu gạch nối (dashes). Những chiếc trống có thể được sử dụng để chuyên tải những thông điệp vượt quá ngoài những khoảng cách về giọng nói con người - một phương pháp được biết như là nói chuyện bằng trống (drum talk). Trong một số trường hợp, những phương pháp giao tiếp như vậy phục vụ chức năng của việc lưu giữ một bí mật thông điệp tránh khỏi bị không thạo. Đây cũng thường là trường hợp với lời nói *whistle speech*, một hình thức của việc giao tiếp trong đó việc huýt sáo thay thế cho lời nói bình thường, thường được sử dụng cho giao tiếp vượt quá những khoảng cách.

II. NGÔN NGỮ HỌC

1 Giới thiệu.

Ngôn ngữ học (linguistics), khoa học nghiên cứu về ngôn ngữ. Nó vây quanh việc miêu tả về các ngôn ngữ, việc nghiên cứu về gốc gác của chúng, và việc phân tích cách thức trẻ con thu nhận ngôn ngữ như thế nào và người ta học các ngôn ngữ khác hơn là học chính tiếng mẹ đẻ của mình ra làm sao. Ngôn ngữ học cũng quan tâm đến những mối quan hệ giữa các ngôn ngữ và đến cách thức và con đường mà các ngôn ngữ biến đổi qua thời gian. Các nhà ngôn ngữ học có thể nghiên cứu ngôn ngữ như một quá trình tư duy và tìm kiếm một lý thuyết nhằm tính toán cho năng lực phổ quát của con người để sản sinh và hiểu ngôn ngữ. Một số nhà ngôn ngữ học khảo sát ngôn ngữ với một ngữ cảnh văn hóa. Bằng việc quan sát những cuộc trò chuyện, họ cố gắng xác định một cá nhân cần cái gì đối với sự hiểu biết để nói năng một cách phù hợp trong những sự thiết lập khác nhau, chẳng hạn như vị trí công việc, giữa những người bạn, hoặc giữa gia đình. Các nhà ngôn ngữ học khác nhấn mạnh đến cái gì sẽ xảy ra khi những người nói từ những ngôn ngữ và những nền văn hóa khác nhau tương tác với nhau. Nhiều nhà ngôn ngữ học có thể cũng tập trung vào cách thức làm thế nào để giúp đỡ những người học ngôn ngữ khác, trong khi sử dụng cái mà họ biết về ngôn ngữ thứ nhất của người học và về ngôn ngữ đang được thu nhận.

2. Ngôn ngữ học miêu tả và so sánh.

Mặc dù có nhiều cách thức và con đường của việc nghiên cứu ngôn ngữ, nhưng đa số các cách tiếp cận đều thuộc về một trong số hai ngành chính của ngôn ngữ học: ngôn ngữ học miêu tả và ngôn ngữ học so sánh.

a. Ngôn ngữ học miêu tả.

Ngôn ngữ học miêu tả (descriptive linguistics) là sự nghiên cứu và phân tích về ngôn ngữ nói (spoken language). Các kỹ thuật của ngôn ngữ học miêu tả từng được nhà nhân chủng học người Đức Franz Boas và nhà ngôn ngữ học kiêm nhà nhân chủng học người Mỹ Edward Sapir nghĩ ra vào đầu những năm 1900 để ghi và phân tích các ngôn ngữ Anh-điêng Mỹ. Ngôn ngữ học miêu tả bắt đầu bằng cái mà nhà ngôn ngữ học nghe những người bản nói. Bằng việc nghe những người nói bản ngữ, nhà ngôn ngữ học thu thập một chỉnh thể về dữ liệu và phân tích nó để xác định những âm thanh khu biệt, được gọi là các *âm vị* (phonemes). Các âm vị riêng lẻ, chẳng hạn như /p/ và /b/, được xác lập trên những nền tảng là sự thay thế của một âm vị này cho một âm vị khác làm thay đổi ý nghĩa của một từ. Sau khi xác định toàn bộ danh mục kiểm kê về các âm tố trong

một ngôn ngữ, nhà ngôn ngữ học xem xét cách thức những âm thanh này kết hợp để tạo ra các *hình vị* (morphemes), hoặc những đơn vị âm thanh mang nghĩa, chẳng hạn như các từ *push* và *bush* như thế nào. Các hình vị có thể là những từ riêng lẻ chẳng hạn như *push*; các từ gốc từ (root words) chẳng hạn như *berry* trong *blueberry*; hoặc những tiền tố (prefixes) (*pre-* trong *preview*) và những hậu tố (suffixes) (*-ness* trong *openness*).

Bước tiếp theo của nhà ngôn ngữ học là xem xét các hình vị kết hợp thành câu như thế nào, trong khi vừa tuân theo cả nghĩa từ điển của âm vị lẫn các quy tắc ngữ pháp của câu. Trong câu *She pushed the bush*, hình vị *she*, một đại từ, là chủ ngữ; *push*, một ngoại động từ, là vị ngữ; *the*, một mạo từ xác định, là định tố; và *bush*, một danh từ, là bổ ngữ. Việc hiểu biết về chức năng của các hình vị trong câu cho phép nhà ngôn ngữ học miêu tả ngữ pháp của một ngôn ngữ. Những thủ tục khoa học về *âm vị học* (phonemics) (cách xác định các âm vị), *hình thái học* (khám phá ra các hình vị), và *cú pháp học* (miêu tả trật tự của các hình vị và chức năng của chúng) cung cấp cho những nhà ngôn ngữ học miêu tả có một cách thức để mô tả các ngữ pháp của nhiều ngôn ngữ chưa bao giờ được mô tả hoặc phân tích trước đó. Với cách thức này, họ có thể bắt đầu nghiên cứu và hiểu các ngôn ngữ đó.

b. Ngôn ngữ học so sánh.

Ngôn ngữ học so sánh (comparative linguistics) là sự nghiên cứu và phân tích, bằng những phương tiện của các bản ghi chép chữ viết, về các nguồn gốc và tính quan hệ của những ngôn ngữ khác nhau. Vào năm 1786, Sir William Jones, một học giả người Anh, đã khẳng định rằng tiếng Sanskrit, tiếng Hy Lạp, và tiếng La-tinh có quan hệ lẫn nhau và đã thừa kế từ một nguồn gốc chung. Ông ta đặt sự khẳng định này trên cơ sở của những nét tương đồng về các âm thanh và ý nghĩa giữa ba ngôn ngữ. Ví dụ, từ Sanskrit *bhratar* cho “anh/em trai” giống với từ La-tinh *frater*, từ Hy Lạp *phrater* (và từ tiếng Anh *brother*).

Các học giả khác tiếp tục so sánh tiếng Băng Đảo với các ngôn ngữ Scandinavian, và các ngôn ngữ Romance với tiếng Sanskrit, tiếng Hy Lạp và tiếng La-tinh. Những sự tương ứng giữa các ngôn ngữ, được biết như là những mối quan hệ cội nguồn, đã đạt được sự biểu hiện trong cái mà các nhà ngôn ngữ học quy chiếu như là các hình cây ngữ hệ, hoặc *hình cây gia phả* (family trees). Những hình cây gia phả được các nhà ngôn ngữ học xác lập bao gồm hình cây gia phả ngôn ngữ Ấn-Âu liên quan đến tiếng Sanskrit, tiếng Hy Lạp, La-tinh, tiếng Đức, tiếng Anh, và các ngôn ngữ châu Âu và châu Á khác; hình cây gia phả ngôn ngữ Algonquian, liên quan đến tiếng Fox, tiếng Cree, tiếng Menomini, tiếng Ojibwa, và những ngôn ngữ bản địa Bắc Mỹ; và hình cây gia phả ngôn ngữ Bantu, liên quan đến tiếng Swahili, tiếng Xhosa, tiếng Zulu, tiếng Kikuyu, và các ngôn ngữ châu Phi khác.

Các nhà ngôn ngữ học so sánh cũng tìm kiếm những nét tương đồng trong cách thức mà các từ được cấu tạo ở những ngôn ngữ khác nhau. Tiếng Anh và tiếng La-tinh, ví dụ, biến đổi mẫu hình thức của một từ để biểu thị những ý nghĩa khác nhau, như khi động từ tiếng Anh *go* biến đổi thành *went* và *gone* để biểu thị một hoạt động đã qua. Tiếng Trung Quốc, mặt khác, không có những hình thức bị biến cách như vậy; động từ bảo lưu như nhau trong khi những từ khác chỉ định thời gian (như trong “go store tomorrow”). Trong tiếng Swahili, các tiền tố, các hậu tố, và các trung tố (những sự thêm vào thân từ) kết hợp với một từ gốc từ để làm thay đổi ý nghĩa của nó. Ví dụ, một từ đơn giản có thể biểu thị khi cái gì đó được hoàn thành, nhờ ai, tới ai, và theo phương thức nào.

Một số nhà ngôn ngữ học so sánh phục nguyên những ngôn ngữ tổ tiên mang tính giả thuyết được biết như là các *ngôn ngữ tiền thân* (proto-languages), họ sử dụng chúng để chứng minh tính quan hệ giữa các ngôn ngữ hiện thời. Tuy nhiên, một ngôn ngữ tiền thân không phải có chủ định miêu tả một ngôn ngữ thực tế, và không phải biểu hiện lời nói của các tổ tiên của những người nói các ngôn ngữ hiện đại. Mĩa mai thay, có một vài nhóm đã sử dụng sai lầm những sự phục nguyên như thế theo một nỗ lực nhằm chứng minh quê hương tổ tiên của một dòng tộc.

Các nhà ngôn ngữ học so sánh đã đề xuất rằng những từ cơ bản nhất định trong một ngôn ngữ không thay đổi qua thời gian, bởi người ta ngại và miễn cưỡng giới thiệu các từ mới cho những từ đã thành hằng số chẳng hạn như “tay”, “mắt”, hoặc “mẹ”. Những từ này được gọi bằng thuật ngữ là *tự do văn hóa* (culture free). Bằng việc so sánh các danh sách của những từ tự do văn hóa trong các ngôn ngữ có cùng ngữ hệ, các nhà ngôn ngữ học có thể chiết xuất ra phần trăm của những từ có quan hệ và sử dụng một công thức để hình dung khi những ngôn ngữ này tách khỏi nhau.

Vào những năm 1960, các nhà so sánh đã không còn thỏa mãn với việc nhấn mạnh đến các nguồn gốc, những sự di trú, và phương pháp hình cây gia phả. Họ thách thức với tư cách là một khái niệm phi thực tế ở chỗ một ngôn ngữ sớm hơn có thể còn lưu lại biệt lập một cách đầy đủ cho những ngôn ngữ khác sẽ được phái sinh từ nó trong cả một giai đoạn thời gian. Ngày nay, các nhà so sánh tìm kiếm để hiểu tính thực tiễn có phức tạp hơn về lịch sử ngôn ngữ, trong việc tính đến tiếp xúc ngôn ngữ. Họ đã đề cập đến các đặc trưng phổ quát của ngôn ngữ và đến những sự so sánh về các ngữ pháp và các cấu trúc.

3. Các lĩnh vực của ngôn ngữ học.

Lĩnh vực ngôn ngữ học vừa mượn và vừa cho các lý thuyết lẫn các phương pháp của riêng mình đối với nhiều bộ môn khoa học khác. Nhiều lĩnh vực hẹp của ngôn ngữ học đã mở rộng sự hiểu của chúng ta về các ngôn ngữ. Các lý thuyết và những phương pháp ngôn ngữ học cũng được sử dụng trong nhiều lĩnh vực nghiên cứu khác. Những sự quan tâm gối chồng lên nhau này đã dẫn tới sự hình thành của một số lĩnh vực-bộ môn đan chéo nhau.

a. Ngôn ngữ học xã hội.

Ngôn ngữ học xã hội (sociolinguistics) là sự nghiên cứu về các mô hình (patterns) và những sự biến đổi về ngôn ngữ cùng với một xã hội hoặc một cộng đồng. Nó tập trung vào cách thức người ta sử dụng ngôn ngữ để biểu thị lớp xã hội, tình trạng nhóm, giới, hoặc tộc người như thế nào, và nó xem xét họ thiết lập những sự chọn lựa về hình thức ngôn ngữ mà họ sử dụng ra làm sao. Nó cũng khảo sát cách thức mà người ta sử dụng ngôn ngữ để điều chỉnh vai trò của mình trong xã hội và để đạt được những vị thế về sức mạnh. Ví dụ, những nghiên cứu ngôn ngữ học xã hội đã nhận ra rằng cách mà một người New York phát âm âm vị /r/ trong một biểu thức chẳng hạn như *fourth floor* có thể chỉ định tầng lớp xã hội của người này. Theo một nghiên cứu, những người mong mỗi chuyển từ giai cấp trung lưu thấp hơn lên giai cấp trung lưu cao hơn thường thêm thanh thế để việc phát âm /r/. Thậm chí đôi khi họ còn overcorrect lời nói của mình, trong khi phát âm /r/ ở nơi mà họ không thể bắt chước một ai đó.

Một số nhà ngôn ngữ học xã hội tin rằng việc phân tích những biến đổi như thế với tư cách là việc sử dụng của một âm vị cụ thể có thể dự đoán phương hướng của sự biến đổi ngôn ngữ. Sự biến đổi, họ nói, vận động hướng đến sự thay đổi được liên tưởng với sức mạnh, uy tín, hoặc phẩm chất khác có giá trị xã hội cao. Nhiều nhà ngôn ngữ học xã hội khác nhấn mạnh đến cái gì xảy ra khi những người nói của những ngôn ngữ khác nhau tương tác. Cách tiếp cận này tới sự thay đổi ngôn ngữ nhấn mạnh cách thức mà các ngôn ngữ pha trộn hơn là phương hướng của sự thay đổi cùng với một cộng đồng. Mục đích của ngôn ngữ học xã hội là để hiểu *ngữ năng giao tiếp* (communicative competence) – cái mà người ta cần phải biết để sử dụng ngôn ngữ phù hợp đối với một bối cảnh xã hội cho sẵn.

b. Ngôn ngữ học tâm lý.

Ngôn ngữ học tâm lý (psycholinguistics) hòa trộn các lĩnh vực tâm lý học và ngôn ngữ học để nghiên cứu người ta quy trình lẫn xử lý ngôn ngữ như thế nào và việc sử dụng ngôn ngữ có liên quan đến việc nêu nền tảng các quá trình tinh thần ra làm sao. Những nghiên cứu về việc thụ đắc ngôn ngữ của trẻ con và việc thụ đắc ngôn ngữ thứ hai là những nghiên cứu mang tính chất ngôn ngữ học tâm lý về bản chất. Các nhà ngôn ngữ học tâm lý làm việc để phát triển những mô hình cho ngôn ngữ được quy trình và được hiểu như thế nào, trong khi sử dụng bằng chứng từ những sự nghiên cứu về cái gì xảy ra khi những quy trình này đi lệch hướng.

Họ cũng nghiên cứu những sự rối loạn ngôn ngữ (language disorders) chẳng hạn như chứng *aphasia* (sự suy giảm khả năng sử dụng hoặc hiểu rõ các từ) và chứng *dyslexia* (sự suy giảm khả năng tạo ra ngôn ngữ viết).

c. Ngôn ngữ học điện toán.

Ngôn ngữ học điện toán (computational linguistics) liên quan đến việc sử dụng các máy tính để biên tập dữ liệu ngôn ngữ, phân tích các ngôn ngữ, dịch từ một ngôn ngữ này sang ngôn ngữ khác, cũng như phát triển và kiểm tra những mô hình của việc xử lý ngôn ngữ. Các nhà ngôn ngữ học sử dụng các máy tính và những mẫu lớn của ngôn ngữ thực tế để phân tích tính quan hệ và cấu trúc của các ngôn ngữ và để tìm kiếm những mô hình và những nét giống nhau. Các máy tính cũng trợ giúp trong nghiên cứu phong cách học, phục hồi thông tin, các hình thái đa dạng của việc phân tích văn bản, lần xây dựng các từ điển và các mục lục. Việc áp dụng máy tính vào các nghiên cứu ngôn ngữ đã để lại kết quả trong những hệ thống dịch máy lẫn máy móc nhận diện và sản sinh lời nói lẫn văn bản. Những máy móc như vậy làm dễ dàng giao tiếp với con người, kể cả những người bị suy giảm về phương diện tri nhận hoặc phương diện ngôn ngữ.

d. Ngôn ngữ học ứng dụng.

Ngôn ngữ học ứng dụng (applied linguistics) áp dụng các lý thuyết và phương pháp ngôn ngữ học vào việc dạy và vào việc nghiên cứu về việc học một ngôn ngữ thứ hai. Các nhà ngôn ngữ học xem xét các lỗi mà người ta làm ra trong khi họ học ngôn ngữ khác và những chiến lược của họ cho việc giao tiếp trong ngôn ngữ mới tại những mức độ khác nhau của ngữ năng. Trong khi tìm kiếm để hiểu cái gì xảy ra trong tư duy của người học, các nhà ngôn ngữ học ứng dụng đoán nhận rằng động lực, thái độ, kiểu học, và tính cá nhân ảnh hưởng đến một người học ngôn ngữ khác tốt ra sao.

e. Ngôn ngữ học nhân chủng.

Ngôn ngữ học nhân chủng (anthropological linguistics), cũng còn được biết như là nhân chủng học ngôn ngữ (linguistic anthropology), sử dụng những cách tiếp cận ngôn ngữ học để phân tích văn hóa. Các nhà ngôn ngữ học nhân chủng khảo sát mối quan hệ giữa một nền văn hóa và ngôn ngữ của nó, cách thức các nền văn hóa và các ngôn ngữ đã thay đổi qua thời gian, cả các nền văn hóa lẫn các ngôn ngữ khác nhau có quan hệ với nhau như thế nào. Ví dụ, cách sử dụng tiếng Anh hiện tại về gia đình và những tên gọi cho sản xuất hiện vào cuối thế kỷ thứ 13 và đầu thế kỷ thứ 14 khi các đạo luật quan tâm đến việc đăng ký, sự chiếm hữu, và sự thừa kế về tài sản được thay đổi.

f. Ngôn ngữ học triết học.

Ngôn ngữ học triết học (philosophical linguistics) khảo sát triết học về ngôn ngữ. Các nhà triết học về ngôn ngữ tìm kiếm những nguyên lý và những xu hướng ngữ pháp mà tất cả các ngôn ngữ của con người chia sẻ. Trong số những sự quan tâm của các nhà triết học ngôn ngữ là phạm vi về những sự kết hợp trật tự từ có thể có xuyên suốt thế giới. Một kết quả tìm kiếm là ở chỗ 95% các ngôn ngữ của thế giới sử dụng trật tự chủ ngữ-động từ-bổ ngữ (SVO) như tiếng Anh sử dụng (*She pushed the bush*). Chỉ 5% sử dụng trật tự chủ ngữ-bổ ngữ-động từ (SOV) hoặc trật tự động từ-chủ ngữ-bổ ngữ (VSO).

g. Ngôn ngữ học thần kinh.

Ngôn ngữ học thần kinh (neurolinguistics) là sự nghiên cứu về ngôn ngữ được xử lý và được biểu hiện trong bộ não như thế nào. Các nhà ngôn ngữ học thần kinh tìm kiếm để xác định những bộ phận của bộ não có liên quan tới sự sản sinh và hiểu ngôn ngữ cũng như để xác định những thành phần của ngôn ngữ (các âm vị, các hình vị, và cấu trúc hoặc cú pháp) được cất giữ ở đâu. Trong khi làm như vậy, họ thiết lập cách sử dụng về các kỹ thuật để phân tích cấu trúc của não bộ và các hiệu ứng của thiệt hại não về ngôn ngữ.