

XÁC SUẤT THỐNG KÊ TRONG NGHIÊN CỨU NGỮ DỤNG HỌC

Hà Cẩm Tâm, M.A

Khoa Anh Đại học Ngoại ngữ, ĐHQG Hà Nội

Việc sử dụng xác suất và thống kê trong nghiên cứu ngôn ngữ nói chung và ngữ dụng học nói riêng là rất phổ biến trên thế giới. Tuy nhiên ở nước ta, việc sử dụng xác suất và thống kê trong nghiên cứu ngữ dụng học vẫn còn là một điều khá mới, vì vẫn còn ít người sử dụng nó. Phần lớn các nghiên cứu mới chỉ dừng lại ở chỗ tính phần trăm và sau đó kết luận, ít công trình sử dụng các phép toán thống kê để kết luận. Báo cáo sau đây muốn giới thiệu tóm tắt về một nghiên cứu ngữ dụng học đã dùng một số phép toán thống kê để phân tích số liệu. Trong báo cáo sau đây chúng tôi xin dừng lại ở phần thông báo về kết quả của phép thống kê được sử dụng để xác định các tình huống cho nghiên cứu ngữ dụng học về lời cầu khiến của người Việt học tiếng Anh và người Úc bản ngữ.

1. ĐẶT VẤN ĐỀ

Trong giảng dạy ngoại ngữ từ những năm 80 trở lại đây người ta chú ý nhiều hơn đến sự liên quan của việc sử dụng ngôn ngữ và các tiêu chuẩn xã hội, văn hoá. Nhiều nghiên cứu đã cho thấy rằng khi học một ngoại ngữ, ngoài việc nắm vững từ vựng, ngữ pháp còn phải nắm được một số những nguyên tắc chi phối việc lựa chọn ngôn ngữ để có thể ứng xử một cách phù hợp. Chẳng hạn khi gặp một người nào đó lần đầu tiên thì trong tiếng Anh chúng ta phải dùng từ chức danh là Mr. hay Mrs. và họ của người ta, nhưng sau đó khi đã biết nhau và trở nên cởi mở với nhau rồi thì người ta có thể yêu cầu mình chỉ gọi người ta bằng tên gọi mà không cần có từ chỉ chức danh đi trước. Chẳng hạn như nếu tên

người đó là John Smith, hay Brian Adam thì gọi là John, hay Brian, chứ không cần gọi là Mr. Smith, hay Mr. Adam. Tương tự như vậy, trong tiếng Việt người học ngoại ngữ phải biết chọn đại từ cho phù hợp, khi chúng ta nói chuyện với một bà già khoảng 70, 80 tuổi thì đại từ phù hợp ít nhất phải là *bà*, nếu không thì phải là *cụ*, chứ không thể là *chị* hoặc *cô*, càng không thể dùng đại từ *em*. Hoặc khi yêu cầu một người lạ mở giúp mình cửa sổ trên tàu hoả thì người Anh phải dùng cấu trúc câu "Would you mind.... " (Nếu không phiền/ Cảm phiền bác) hoặc là "I was wondering if you could.... " (tôi không biết là bác có thể/ liệu bác có thể) chứ không thể nói một cách giản đơn là "Could you... " (bác có thể...) hoặc là "Can you.... " (bác giúp hộ...)

Theo hướng trên, người ta đã tiến hành nhiều nghiên cứu để tìm hiểu xem người Nhật học tiếng Anh, hay người Trung Quốc học tiếng Anh thường mắc những lỗi gì trong việc lựa chọn các hình thức ngôn ngữ của tiếng nước ngoài khi giao tiếp, và những lỗi đó do đâu mà có. Các nghiên cứu trên thế giới của Blum-Kulka et al. (1989), Trosborg (1995), Tanaka (1988) v.v. đã nhận thấy rằng so với người bản ngữ người học ngoại ngữ thường không tìm được hình thức ngôn ngữ phù hợp trong nhiều tình huống giao tiếp do không nắm vững được một số quy tắc văn hoá quy định sự lựa chọn, hoặc người học lại sử dụng quy tắc dụng học của tiếng mẹ đẻ vào ngoại ngữ. Do đó tạo nên những câu nói nhiều khi không lịch sự bằng hoặc lại quá kiêu cách, quá lịch sự so với người bản ngữ.

Để phân nào giúp người Việt Nam học tiếng Anh nói được tiếng Anh đúng hơn, chúng tôi đã tiến hành một nghiên cứu về ngữ dụng học. Cụ thể là chúng tôi đã nghiên cứu việc thực hiện một hành động lời nói, lời yêu cầu (request), của người Việt Nam học tiếng Anh và người Úc trong một số tình huống nhất định. Mục đích là tìm hiểu xem người Việt nam có bị sai lệch so với người bản ngữ không, và nếu có thì sai lệch nhiều hay ít, có thể tìm hiểu nguyên nhân của sự sai lệch đó không? Tuy nhiên trước khi tìm hiểu những sai lệch về hình thức ngôn ngữ nói trên, chúng ta cần phải tìm hiểu xem trong cùng một tình huống giao tiếp, người Việt và người Úc có khác nhau trong cách đánh giá các thông số xã hội không.

2. THIẾT KẾ NGHIÊN CỨU

Với mục tiêu trên chúng tôi đã thiết kế nghiên cứu của mình như sau.

Trước tiên, chúng tôi xây dựng một ngân hàng các tình huống mà trong đó đối tượng nghiên cứu phải đưa các lời yêu cầu. Theo Brown & Levinson (1987) thì mặc dù có một số yếu tố được coi là có tác động đến sự lựa chọn hình thức phù hợp của lời yêu cầu nhưng ba yếu tố được coi là quan trọng nhất và có những ảnh hưởng rõ nhất là: (1) quan hệ xã hội giữa người nói và người nghe (cụ thể là thân hay không thân, quen biết hay không quen biết v.v.) được viết tắt là D; (2) địa vị tương đối của hai người đối thoại với nhau (thể hiện ở chỗ người nào có quyền ra lệnh cho người nào v.v.) được viết tắt là P; (3) mức độ áp đặt của lời yêu cầu đối với người nghe được viết tắt là R. Các tác giả này cũng chỉ ra rằng các yếu tố này phụ thuộc vào hoàn cảnh và văn hoá. Điều này có nghĩa là cùng một hoàn cảnh giao tiếp các nền văn hoá khác nhau sẽ có những đánh giá khác nhau về các thông số xã hội nói trên, có nghĩa là D, P, R có thể có những giá trị khác nhau. Do đó sẽ dẫn đến những sự lựa chọn khác nhau hoặc những quy ước khác nhau về các hình thức ngôn ngữ phù hợp.

Trên cơ sở lý thuyết của Brown & Levinson (1987) các tình huống được xây dựng với những biến là các giá trị tương đối của các yếu tố xã hội đã nêu trên. Cũng theo hai tác giả này thì các biến này là các biến phụ thuộc, do vậy khi P và R được giữ nguyên và có giá trị nhỏ theo đánh giá của người nói, nói cách khác tức là khi địa vị của người nói và người nghe phần nào bình đẳng và mức độ áp đặt không lớn thì chỉ còn quan hệ xã hội là yếu tố quyết định cho việc lựa chọn hình thức ngôn ngữ phù hợp (Brown & Levinson, 1987, p.80).

Theo đó 32 tình huống đã được xây dựng bao gồm các các giá trị khác nhau của hai trong ba biến nhắc đến ở trên, đó là (1) quan hệ tương đối về vị thế (P), và (2) khoảng cách tương đối về xã hội (D). Trong nghiên cứu này do hạn chế về thời gian và khuôn khổ của luận văn nên mức độ áp đặt của yêu cầu (R: ranking of imposition) được không chế ở mức nhỏ. Trong nghiên cứu này mối quan hệ vị thế P của hai người giao tiếp có các giá trị là +P, =P và -P. Ba giá trị đó phản ánh ba giá trị

về vị thế của người nói và người nghe, lần lượt là: (1) người nói có vị thế cao hơn người nghe, có quyền ra lệnh cho người nghe, chẳng hạn như xếp và thư ký; (2) người nói và người nghe ngang bằng nhau về vị thế, chẳng hạn như hai người đều là nhân viên của một văn phòng, có vị trí công tác tương đương nhau; (3) người nói có vị thế thấp hơn người nghe, chẳng hạn như nhân viên và giám đốc.

Tiêu chí thứ hai là khoảng cách xã hội, được hiểu là sự thể hiện mức độ gần gũi, thân mật giữa hai người giao tiếp. Theo đó nếu hai người là người trong gia đình, hay là bạn cùng làm trong một văn phòng quen biết nhau lâu sẽ được coi là có khoảng cách xã hội nhỏ và được ký hiệu là -D. Ngược lại nếu hai người không quen biết nhau thì khoảng cách xã hội lớn, không gần gũi nhau, và được ký hiệu là +D.

Với các giá trị trên của các thông số ta cần có 6 tình huống, cụ thể là:

[+P, +D] ; [=P, +D] ; [-P, +D] ; [+P, -D] ; [=P, -D] ; [-P, -D]

Để xác định được 6 tình huống này, chúng tôi đã đưa 32 tình huống đi khảo sát. Điều trước tiên phải khảo sát là để nắm vững được giá trị của các thông số xã hội trong các tình huống giao tiếp thông qua sự đánh giá của đối tượng nghiên cứu. Các thông số xã hội trong các tình huống lần lượt được đưa ra kiểm tra, đối tượng nghiên cứu đánh giá chúng theo hệ thống "likert scale" 5 cấp độ. Việc đánh giá các thông số xã hội của 32 tình huống được làm hai lần trên cùng một đối tượng, hai lần cách nhau là 2 tuần. Sau đó phần mềm thống kê SPSS được sử dụng để phân tích số liệu. Cụ thể là phép thử paired t-test đã được sử dụng. Theo Coakes & Steed (1996, tr. 58-59) thì:

Phép thử t-test lặp, còn được gọi là t-test cặp đôi phụ thuộc, được dùng khi người ta dùng số liệu chỉ thu thập từ một nhóm đối tượng nghiên cứu...Số liệu được thu thập từ một nhóm đối tượng nghiên cứu còn được gọi là số liệu trong nhóm (within-subjects), bởi vì nhóm đối tượng đó đã thực hiện việc cung cấp số liệu hai lần. Những nghiên cứu sử dụng thiết kế "pretest-post test" thường được phân tích bằng phép thử t-test lặp. Mục đích là để xác định xem sự khác nhau giữa hai giá trị "means" của hai lần lấy kết quả có thật sự khác nhau không.

Vì mục đích của phép thử này là để tìm ra những tình huống trong đó các giá trị của các thông số xã hội được đánh giá một cách nhất quán, cho nên những kết quả khẳng định giả thuyết không (null hypothesis)

sẽ được chấp nhận cho nghiên cứu này. Như vậy, giả thuyết cần được chấp nhận ở đây là "giá trị của các thông số xã hội P, D, R sau hai lần kiểm nghiệm không khác nhau." Điều đó cũng có nghĩa là, người ta trong một chừng mực nào đó phải giống nhau về cách đánh giá các thông số xã hội được sử dụng trong các tình huống nghiên cứu. Ngưỡng để phản bác giả thuyết không là $p < 0,05$, theo đó phép thử t-test hai đầu sẽ được sử dụng, bởi nếu có sự khác nhau, thì người ta chưa thể khẳng định nó về hướng nào. Sau đây là một ví dụ về một trong những tình huống được khảo sát theo hệ thống likert scale năm cấp độ:

22. You have your own car. Your car is not as reliable as your elder brother's car. You want to go out with your friends this Sunday. You know your brother is not using his car this Sunday. You want to borrow it.

	1	2	3	4	5
How much authority or right do you think the speaker has in making the request?	none		some		a great deal
How well acquainted are the speaker and the hearer?	not at all		a little bit		very well
How large is the imposition of the request on the hearer?	Very small		moderate		very large

3. KẾT QUẢ NGHIÊN CỨU

Sau khi thực hiện phép thử chúng tôi đã thu được kết quả. Trước tiên, như đã trình bày ở trên chúng tôi loại bỏ tất cả những tình huống có giá trị $p < 0,05$, bởi chúng tôi đang tìm những tình huống mà đối tượng nghiên cứu nhất quán trong các đánh giá các giá trị xã hội. Như vậy là những kết quả có giá trị $p \geq 0,05$ sẽ được giữ lại. Bởi giá trị của p càng cao thì độ nhất quán càng cao do đó độ tin cậy càng tăng. Dưới đây là một số kết quả khảo sát của đối tượng nghiên cứu là người Úc:

The results of the T-test of the P, D and R (Australian data)

Sit	P					D					R				
	mean	sd	t-val.	df	p	mean	sd	t-val.	df	p	mean	sd	t-val.	df	p
15	3.02	1.23	-.23	91	.820	1.05	0.23	-1.16	91	.250	1.89	1.03	.51	91	.615
	3.05	1.25				1.10	0.30				1.83	1.04			
16	3.94	1.04	.86	91	.391	1.25	0.66	.29	91	.770	1.97	1.08	-.50	91	.618
	3.84	1.01				1.23	0.47				2.02	1.00			
17	2.19	1.03	-.67	91	.502	1.94	0.84	-2.81	91	.006	4.29	0.98	1.42	91	.158
	2.27	1.01				2.24	0.89				4.16	0.96			
18	2.30	1.00	.60	91	.547	3.12	0.74	1.74	91	.085	2.99	1.00	.75	91	.457
	2.24	0.83				2.97	0.73				2.90	0.93			
19	2.36	0.88	.38	91	.705	4.38	0.81	2.58	91	.012	2.65	0.92	.41	91	.682
	2.31	0.96				4.05	0.99				2.61	0.75			
20	2.41	0.96	1.07	91	.285	4.34	0.84	-.67	91	.505	2.86	0.83	1.87	91	.064
	2.29	0.88				4.42	0.99				2.63	0.92			
21	3.35	1.05	-.26	91	.799	3.18	0.85	1.05	91	.294	2.92	0.87	.96	91	.330
	3.38	1.06				3.08	0.84				2.81	0.89			
22	2.89	1.08		91	.320	4.94	0.27	1.69	91	.095	2.76	1.16	.29	91	.772
	3.01	1.23	1.00			4.85	0.51				2.72	1.11			
23	3.03	1.13	.27	91	.787	3.71	0.83	-.12	91	.904	2.98	0.77	1.89	91	.062
	3.00	1.16				3.72	0.84				2.79	0.76			
24	3.08	1.22	.75	91	.457	2.71	0.60	1.40	91	.164	1.90	0.96	-.76	91	.451
	2.99	1.24				2.58	0.79				2.00	1.04			
25	2.46	1.22		91	.210	1.93	1.02	-1.05	91	.295	2.73	1.28	.24	91	.812
	2.63	1.31	1.26			2.06	0.92				2.69	1.28			
26	3.95	0.98	.18	91	.856	3.28	0.84	-.67	91	.502	2.55	0.91	.19	91	.847
	3.92	0.99				3.36	0.76				2.53	0.93			
27	3.04	1.18	-.28	90	.783	2.83	0.81	-1.04	91	.301	2.81	0.89	.96	91	.330
	3.08	1.17				2.92	0.65				2.71	0.91			

So sánh với một số kết quả của đối tượng nghiên cứu là người Việt

The results of the T-test of the P, D and R (Vietnamese data)

Sit	P					D					R				
	mean	sd	t-val.	df	p	mean	sd	t-val.	df	p	mean	sd	t-val.	df	p
15	2.11	1.17	1.19	105	.236	1.71	1.3	.00	105	1.000	2.28	0.96	.78	105	.435
	2.25	0.94				1.71	1.3				2.36	0.85			
16	3.34	1.03	.09	105	.927	2.30	1.11	-.60	105	.552	2.99	1.15	-.26	105	.796
	3.35	0.89				2.23	0.96				2.96	1.00			
17	2.28	1.04	2.35	105	.021	2.89	0.98	.25	105	.802	3.06	1.23	-1.25	105	.213
	2.57	1.03				2.91	0.88				2.90	1.16			
18	2.25	1.14	.09	105	.930	3.33	1.11	1.35	105	.181	2.83	1.24	.58	105	.566
	2.26	1.03				3.48	0.95				2.89	1.14			
19	2.48	1.14	.08	105	.934	3.93	0.96	-1.74	105	.085	2.65	1.09	1.34	105	.184
	2.49	0.95				3.73	1.02				2.80	0.99			
20	2.99	1.08	.17	105	.864	4.49	0.95	-.44	105	.664	2.88	1.20	.00	105	1.000
	3.01	0.89				4.44	0.83				2.88	1.10			
21	2.90	1.13	.97	105	.336	3.20	0.96	.18	105	.861	3.05	1.04	-1.64	105	.103
	3.01	0.99				3.22	0.98				2.90	0.92			
22	2.76	1.14	1.53	105	.130	4.46	0.90	-.89	105	.376	2.98	1.10	-.24	105	.813
	2.94	1.05				4.37	0.98				2.95	1.06			
23	2.65	1.07	-1.67	105	.097	3.51	1.06	.00	105	1.000	2.64	1.16	.84	105	.401
	2.47	0.85				3.51	0.96				2.73	0.88			
24	2.49	1.16	-0.64	105	.524	3.02	0.97	-.27	105	.788	2.57	1.13	.51	105	.609
	2.41	1.01				2.99	0.90				2.63	0.97			
25	2.74	3.25	-1.02	105	.311	3.03	1.17	-.10	105	.922	2.65	1.20	-1.24	105	.219
	2.43	1.06				3.02	1.13				2.53	1.01			
26	3.32	1.02	-.94	105	.347	3.58	0.89	-.20	105	.839	3.27	1.12	-.87	105	.386
	3.23	0.97				3.57	0.94				3.17	1.07			
27	2.49	0.998	-.16	105	.874	3.06	1.06	-.61	105	.542	2.68	1.05	.55	105	.586
	2.47	1.05				2.98	1.01				2.74	0.97			

Chúng tôi đã dà soát các giá trị của P, D và R trên cơ sở so sánh các kết quả của means, standard deviation (sd.), và giá trị của p để xác

định những tình huống có độ tin cậy cao nhất đồng thời thoả mãn các giá trị mà chúng tôi cần cho nghiên cứu của mình. Theo cách trên chúng tôi đã xác định được các tình huống đủ độ tin cậy theo đánh giá của người Úc và các tình huống đủ độ tin cậy theo đánh giá của người Việt. Trong số những tình huống đủ độ tin cậy chúng tôi lại lựa ra những tình huống có độ tin cậy cao nhất trong nhóm và dùng chúng trong phần khảo sát về hình thức ngôn ngữ. Kết quả là chúng tôi đã xác định được 6 tình huống được coi là đủ độ tin cậy nhất đối với cách đánh giá của người Úc và 6 tình huống được coi là đủ độ tin cậy nhất đối với người Việt.

Đúng như một số nhà nghiên cứu đã tuyên bố, trong cùng một tình huống giao tiếp, thì người thuộc các nền văn hoá khác nhau có thể có cách đánh giá khác nhau. Và đây là một trong những nguyên do chính dẫn đến sự ứng xử ngôn ngữ khác nhau. Quan sát hai bảng trên ta có thể thấy sự khác nhau khá xa giữa đối tượng nghiên cứu là người Úc và đối tượng nghiên cứu là người Việt nam. Trong bảng chúng tôi đã tô đậm những tình huống đủ độ tin cậy cho nghiên cứu theo đánh giá của hai nhóm đối tượng. Chẳng hạn như tình huống 16 hai nhóm có những đánh giá khá khác nhau. Người Úc và người Việt cho chúng những giá trị means khác nhau. Trong kết quả khảo sát tình huống này của người Úc P có means là 3,94 và 3,84; D có means là 1,25 và 1,23; và R có means là 1,97 và 2,02. Người Việt lại có P với means bằng 3,34 và 3,35; D với means bằng 2,30 và 2,23; R với means bằng 2,99 và 2,96. Như vậy nếu P của người Úc tiến gần hơn tới +P cực đại, tức là người nói có quyền yêu cầu người nghe thì P của người Việt chỉ nhích hơn giá trị của =P một chút, có nghĩa là người nói có được quyền yêu cầu người nghe nhưng quyền đó không lớn như đối với người Úc. Nếu xem xét kỹ chúng ta sẽ còn thấy nhiều sự khác biệt nữa.

4. KẾT LUẬN

Có thể thấy rõ rằng trong các nền văn hoá khác nhau người ta có những cách nhìn nhận khác nhau về các mối quan hệ. Chính điều này đã làm nên sự khác biệt trong hành động lời nói. Cũng chính vì lý do này mà khi nghiên cứu về ngữ dụng học chúng ta cần phải tiến hành

kiểm tra độ tin cậy của công cụ thu thập số liệu. Để đảm bảo độ tin cậy cho công cụ thu thập số liệu thì cần phải sử dụng các phép kiểm tra của thống kê để đảm bảo rằng những thông tin và kết luận mà chúng ta đưa ra không bị chủ quan, không bị sai lệch.

SÁCH THAM KHẢO

1. Blum-Kulka, S. (1989). Playing it safe: The role of conventionality in indirectness. In S. Blum-Kulka., J. House, & G. Kasper (Eds.), *Cross cultural pragmatics: Requests and apologies* (pp. 37-70). Norwood, N.J.: Ablex.
2. Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
3. Coakes, S. J. & Steed, L. G. (1996). *SPSS for windows: Analysis without anguish*. Singapore: John Willey & Son.
4. TAM, H.C. (1998). *Request by Australian Native Speakers of English and Vietnamese Learners of English - a cross communication study in pragmalinguistics*. M.A thesis, La Trobe university, 1998.