A Study on the Oral Disfluencies Developmental Traits of EFL Students — A Report Based on Canonical Correlation Analysis

Ying Gao[1] & Wanyi Du[1]

[1] School of Software, Dalian University of Technology, Dalian, Liaoning Province, China

Correspondence: Ying Gao, School of Software, Dalian University of Technology, Dalian, Liaoning Province, China. Tel: 86-411-8757-1529. E-mail: gaoying001@126.com

Abstract

This paper traces 9 non-English major EFL students and collects their oral productions in 4 successive oral exams in 2 years. The canonical correlation analysis approach of SPSS is adopted to study the disfluencies developmental traits under the influence of language acquisition development. We find that as language acquisition develops, the total production of difluenices does not decrease correspondingly as we thought, but keeps constant for a period of time. While the proportions of specific disfluencies phenomena change significantly, which features the decrease of pauses and the increase self-repairs. Besides, the grammatical accuracy and language complexity have opposite effects on disfluencies traits. In the first year, disfluencies were displayed mainly as pauses and repetitions since EFL students paid more attention to grammatical accuracy; in the second year, disfluencies featured more self-repairs and less pauses because EFL students transferred their attention to language complexity. We also find language acquisition can only account for partial developmental traits of disfluencies despite of the strong correlations between them, and other factors, such as psychological or social elements, may also take effects.

Keywords: disfluency, language acquisition development, canonical correlation analysis

1. Introduction

For EFL students, one of their purposes is to improve their oral English. But the majority of Chinese EFL students cannot communicate fluently even after several years of English learning, which is far from their expectations. In daily classes, instead of giving specific advices to students directed at their individual oral problems, teachers always tend to offer general and vague instructions, such as "pay attention to the accuracy of languages" or "try to improve your pronunciation", etc. These kinds of instructions benefit students less. Students are not clear about their oral problems, nor do they notice their improvement in oral English. So it is very common for Chinese EFL students to give up oral practices without a sense of achievement after a period of time. This phenomenon has a close relationship with the lack of understanding and researches in oral disfluencies of foreign languages.

Oral disfluencies generally refer to the non-fluent parts in oral productions (Shriberg, 1994). It may also refer to the disjointed or relatively slow oral parts in communications (Starkeweather, 1987). From these definitions, we find disfluencies are displayed not only as broken languages, but also as self-repairs and languages errors, etc. Dollaghan and Campbell (1992) studied the "disfluencies traits" system and classified disfluencies into 4 groups: pauses, repetitions, self-repairs and orphans. This paper accepts the 4 groups of disfluencies and considers disfluencies as the oral outputs which make oral productions disfluent or unnatural. Dollaghan and Campbell (1992) suggested that each group of disfluencies was an independent phenomenon, which reflected a corresponding language learning process. Many scholars have studied disfluencies (Baars, Motley, & MacKay, 1975; Dell & Reich, 1981; Fromkin, 1971; Garrett, 1975; Lee, 1974; Pearl & Bernthal, 1980; Wall & Myers, 1984). Majority of them studied certain aspect of disfluencies traits by analyzing the associations between language proficiency and disfluencies. While majority of these researches studied the disfluencies traits at a specific time rather than the longitudinal developmental changes. Besides, most of researches focused on changes of disfluencies traits under one certain language acquisition phenomenon (such as syntax) (Gordon & Peterson, 1986; Colburn & Mysak, 1982) other than more language phenomena. Thus canonical correlation analysis approach was seldom used in these researches.

In this paper, we will study oral disfluencies developmental traits longitudinally by analyzing its correlations with language acquisition development. More than one language acquisition phenomena will be considered, thus canonical correlation analysis approach will be applied in it.

2. Methodology

Canonical correlation analysis, introduced by Harold Hotelling in 1936, is a way of making sense of cross-covariance matrices. If we have two sets of variables, $x_1....x_n$, and $y_1...y_n$, and there are correlations among the variables. Canonical correlation analysis enables us to find linear combinations of the x's and the y's which have maximum correlation with each other. The linear combinations are called pairs of canonical variates. The coefficients of the pairs of canonical variates show maximum correlations. We need to study several pairs of canonical variates to find out the correlations between these two sets of variables. The second pair of canonical variate is the pair which has the second biggest coefficient and is uncorrelated with the first pair. Then we can use the same way to find the third pair, the fourth and others. When we summarize the correlations of these pairs of canonical variates, we can get nearly all correlations information between the two sets of variables. While one or two pairs of canonical variates are enough to show the correlations in general.

In this experiment, we selected 9 non-English major EFL college freshmen of Dalian University of Technology at random and traced their successive 4 oral English tests productions in 2011-2012. Every student produced a 3-minute speech. In oral tests, students were supposed to draw lots for their topics. Next, we transcribed their tape-recordings into words and labeled the disfluencies signals and language acquisition developmental indicators. Finally, we used canonical correlation analysis approach of SPSS 13.0 to analyze these data to find out the connections between disfluencies phenomena variables and language acquisition development variables.

2.1 Disfluencies Phenomena Variables

According to the "disfluencies traits system" of Dollaghan and Campbell (1992), disfluencies can be divided into 4 categories. So disfluencies phenomena variables include: pause ratios, repetition ratios, self-repair ratios and orphan ratios.

X_1 =pause ratios. Pauses in this paper refer to the intermissions in sentences or between sentences longer than 0.3 seconds (Raupach, 1987).

X_2=repetition ratios. Repetitions refer to the repeated parts taking place in the same sentences and the repeated parts are conjoint (pauses may happen in the middle).

X_3=self-repair ratios. Self-repairs are defined as the error revisions in syntactic frames, lexical structures, tenses or pronunciations.

X_4=orphan ratios. Orphans refer to the intrusion of seemingly unrelated materials to topics.

2.2 Language Acquisition Development Variables

The criteria about spoken language proficiency may vary among different researchers (Galloway 1987; McNamara 1996). Higgs and Clifford (1980) proposed Relative Contribution Model (RCM) and suggested that different factors contribute differently to overall language proficiency at differently levels. In RCM model, vocabulary and pronunciation factors are most important at the beginning levels. At the higher level, fluency and grammar make contributions. At the highest levels, all these four factors and sociolinguistic factor work together for greatest language proficiency. This paper adopts this RCM Model and sets language acquisition development (proficiency) variables from the perspectives of vocabulary, grammatical accuracy, grammatical complexity and pronunciation.

Y_1=simple sentences ratios (the total numbers of simple sentences to total utterance)

Y_2=compound sentences ratios (the total numbers of compound sentences to total utterance)

Y_3=complex sentences ratios (the total numbers of complex sentences to total utterance)

Y_4=non-complete sentences ratios (the total numbers of non-complete sentences to total utterance)

Y_5= type-token ratios (the total number of different words to total number of words)

Y_6=new sematic contents (the utterance of non-repeated complete languages to the total utterance). For example, in a sentence such as "…Yeah, we should… we should care….just as it is, we should care about what we said. It is very important. We can't tell……we can't tell….we can't tell about some personally things". The new sematic contents are "we should care about what we said", "it's very important" and "we can't tell about some personally (personal) things". Their total utterance is 25, while the total utterance of these sentences is 40, so the new sematic content ratio is 62.5%. If EFL students repeat certain opinions simply in oral productions, the repeated

parts cannot be regarded as new semantic contents. New semantic contents ratios reflect the control capabilities of students on topics to some extent.

Y_7=the number of phrases per T-unit

Y_8= error free T-units to the total number of T-units. T-unit is defined as the independent clauses and clauses affiliated with other clauses (Hunt, 1970).

Y_9=the number of clauses per T-unit.

Y_{10}=the average phonological complexity of vocabulary. The phonological complexity is decided by the following way: the word which has fewer than 3 phonemes=1;the word which has 3-4 phonemes=2 and greater than 4 phonemes=3. Using this way, we can give a number to each word and the average number of each oral production can be calculated accordingly (Masterson and Kamhi, 1992).

Y_{11}= phonological accuracy. It equals to the number of words with correct pronunciations to the total number of words.

3. Results

3.1 Disfluencies Productions in these 4 Successive Oral Tests

After labeling the disfluencies signals and calculating, we get the data related to disfluencies traits and show them in Table 1 and Figure 1:

Table 1. Disfluencies and other measures related to disfluencies

	Total number of disfluencies	Total utterance	Total number of words	Mean speech rate	Mean length of utterance
1st Test	165	2487	2460	107.43	15.07
2nd Test	258	3765	3699	151.94	14.59
3rd Test	240	4635	4074	167.02	19.31
4th Test	156	3906	3531	166.7	25.03

Note: mean length of utterance=total utterance/total numbers of disfluencies; the data in Table 1, total utterance and total number of words are produced by 9 students in 27 minutes, while mean rate of speech and mean length of utterance are the average number of 9 students in each oral test.

As shown in Table 1, the total number of disfluencies produced within the same 3 minutes increases from the first test to the second test, keeps nearly constant during the interval between the second test to the third test and finally decreases in the fourth test.
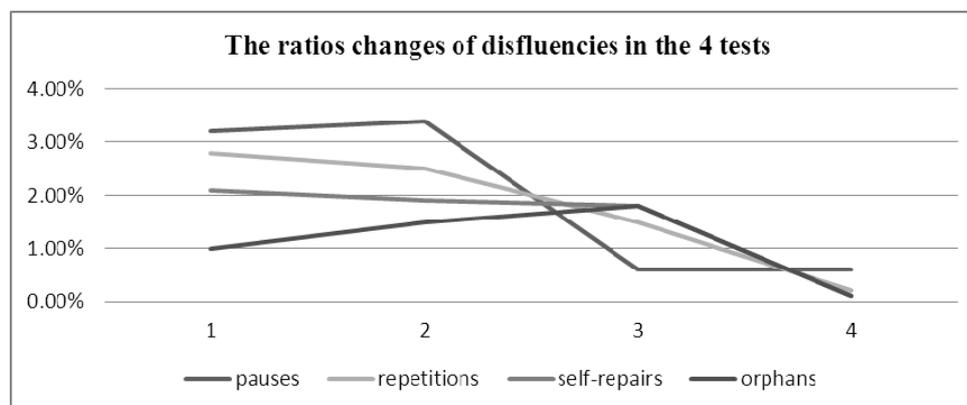

Figure 1. The weights changes of disfluencies phenomena in the 4 tests

Note: The weight (ratio) of each kind of disfluencies= the number of each kind of disfluencies/the total utterance

We also find the obvious ratios changes of specific disfluencies phenomena from Figure 1. The ratios of pauses, repetitions and orphans keep decreasing, while the ratios of self-repairs increase in these 4 tests. In the first and second oral test, pauses and repetitions are the main forms of disfluencies; while self-repairs become the significant disfluencies in place of pause and repetitions in last two tests,

The third tendency we can find (from Table 1) is that the mean speech rates increases apparently while the mean length of utterance increases slightly. With the increase of mean speech rates, the number of utterance per minute increases accordingly and the total number of disfluencies rises too. So the mean length of utterance changes slightly. As for the listeners, little progress has been made in EFL students' first three oral productions. While in the last test, the listeners can notice the difference in language expressions and can feel the improvement of oral English proficiency.

3.2 Findings from Canonical Correlation Analysis

In SPSS, there is no ready menu for canonical correlation analysis, so we write down the following sentences in grammar window (File—New—Syntax):

    INCLUDE 'C: \Program Files\SPSS\Canonical correlation.sps'.

    CANCORR SET1=pause repetition revise orphans /

    SET2=simple compound complex non-complete var9 var10 var11 var12 var13 complexity accuracy/.

When writing down these orders, we have to note that in a CANCORR macroprogram, INCLUDE sentence is used to read macroprogram related to canonical correlation, and the position of macroprogram can vary with different installation catalogues. Besides, INCLUDE and CANCORR these two sentences should be finished with full stops (.). After entering the program above, select menu Run->All and operate this program, then we get the following results about canonical correlation analysis (shown in Table 2 and Table 3):

Table 2. Correlations for Set-1

	X_1	X_2	X_3	X_4
X_1	1.0000	.6656	.2157	.4662
X_2	.6656	1.0000	.3562	.2218
X_3	.2157	.3562	1.0000	-.0097
X_4	.4662	.2218	-.0097	1.0000

Table 3. Correlations for Set-2

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}
Y_1	1.0000	.0496	-.2267	.0169	.0821	-.1823	-.3318	-.3761	-.4589	.2830	-.4569
Y_2	.0496	1.0000	.1680	-.0646	-.2434	.2304	.3344	.3179	.2201	-.0463	.0628
Y_3	-.2267	.1680	1.0000	-.0461	-.3569	.5547	.6596	.6317	.5556	-.1417	.4163
Y_4	.0169	-.0646	-.0461	1.0000	-.0589	-.0234	.0054	-.0336	.0334	.1284	.0928
Y_5	.0821	-.2434	-.3569	-.0589	1.0000	-.4741	-.6477	-.6970	-.7292	.5044	-.5468
Y_6	-.1823	.2304	.5547	-.0234	-.4741	1.0000	.8164	.7045	.7441	.0714	.4797
Y_7	-.3318	.3344	.6596	.0054	-.6477	.8164	1.0000	.9671	.9112	-.2284	.5746
Y_8	-.3761	.3179	.6317	-.0336	-.6970	.7045	.9671	1.0000	.9240	-.3913	.6181
Y_9	-.4589	.2201	.5556	.0334	-.7292	.7441	.9112	.9240	1.0000	-.3745	.7546
Y_{10}	.2830	-.0463	-.1417	.1284	.5044	.0714	-.2284	-.3913	-.3745	1.0000	-.4102
Y_{11}	-.4569	.0628	.4163	.0928	-.5468	.4797	.5746	.6181	.7546	-.4102	1.0000

From the matrix for each set in Table 2 and Table 3, we find the variables inside each set are correlated. For instance, the 4 variables are correlated in set-1 and their coefficients are appropriate, not too big nor too small. So each variable can represent one kind of disfluencies and cannot be replaced by another. If the coefficient

between two variables is too big, it is necessary to consider the combination of these two variables, or just delete one (Tao Zhou, 2010). While although the coefficients between Y_7 and Y_8, Y_7 and Y_9, Y_8 and Y_9 are big enough to consider the combination issue in set-2, we still keep these 3 variables unchanged because there are no real overlaps in contents among these 3 variables. All in all, the variables in these two sets are well selected and are good representatives of certain aspect of each set.

Table 4. Coefficients of Set-1 and Set-2

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}
X_1	.3842	.0587	-.4065	.1342	.0792	-.2068	-.3886	-.4142	-.3783	.1153	-.2386
X_2	.3593	.0858	-.2750	.2093	.0713	.0050	-.1213	-.2078	-.1480	.2516	-.2307
X_3	.2103	.2282	.2705	-.0449	-.3767	.6285	.6092	.4902	.4790	.0783	.1471
X_4	.0769	-.1488	-.5479	-.0102	.2184	-.3318	-.2996	-.2338	-.2270	-.1792	-.1551

From Table 4, we find the direct coefficients between variables of these two sets are not great, except the 2 greater coefficients of self-repairs (X_3) and new semantic content (Y_6), and of self-repairs (X_3) and the number of phrases per T-unit(Y_7) (R=0.6285 and R=0.6092). While in Table 5, the first to the third canonical coefficients (R=0.922, R=0.762 and R=0.718) are bigger than any simple coefficients in Table 4. This proves that the effects of comprehensive canonical correlations are greater than that of the simple correlations among variables. That is, the language acquisition variables as a whole have stronger impact on the disfluencies developmental traits.

Table 5. Canonical Coefficients and Test that remaining correlations are zero:

	Canonical Correlations			Test that remaining correlations are zero:			
					Wilk's	Chi-SQ	DF	Sig.
1	.922		1		.026	98.718	44.000	.000
2	.762		2		.172	47.507	30.000	.022
3	.718		3		.411	24.004	18.000	.155
4	.390		4		.848	4.448	8.000	.815

The significance test results in Table 5 show that when a=0.05, the first and the second canonical coefficients are significant, while the third and the fourth ones are not. So the correlations between two sets of variables are reflected by the correlations of these two pairs of canonical variates. In order to eliminate the influence of different dimensions and units of raw variables, we adopt standardized canonical coefficients (in Table 6) and set up the linear models in Table 7:

Table 6. Standardized Canonical Coefficients for these 2 sets:

Standardized Canonical Coefficients for Set-1					Standardized Canonical Coefficients for Set-2				
	1	2	3	4		1	2	3	4
X_1	.603	-.881	-.575	.867	Y_1	-.385	-.716	.548	-.054
X_2	-.340	-.320	.740	-1.104	Y_2	.120	-.276	-.129	-.072
X_3	-.890	.048	-.070	.597	Y_3	.121	-.101	-.925	-.673
X_4	.036	.561	.948	.313	Y_4	.152	-.179	.121	-.282
					Y_5	-.246	.280	1.036	-.543
					Y_6	.364	-.216	-.317	1.023
					Y_7	-2.852	-.386	.779	.147
					Y_8	2.098	-1.396	-.448	1.065
					Y_9	.927	-.324	1.316	-2.407
					Y_{10}	.050	-.008	-.507	-.542
					Y_{11}	.285	-.126	-.286	.573

Table 7. Canonical correlation models:

No.	Canonical correlation models
1.	$U_1=0.603X_1-0.340X_2-0.890X_3+0.036X_4$
	$V_1=-0.385Y_1+0.120Y_2+0.121Y_3+0.152Y_4-0.246Y_5+0.364Y_6-2.852Y_7+2.098Y_8-0.927Y_9+0.05Y_{10}+0.285Y_{11}$
2.	$U_2=-0.881X_1-0.320X_2+0.048X_3+0.561X_4$,
	$V_2=-0.716Y_1-0.276Y_2-0.1.1Y_3-0.179Y_4+0.280Y_5-0.216Y_6-0.386Y_7-1.396Y_8-0.324Y_9-0.008Y_{10}-0.126Y_{11}$.

In the first pair of canonical variates, U_1 represents the first variate of disfluencies; V_1 represents the first variate of language acquisition (proficiency). The coefficients of X_1 and X_3 are 0.603 and -0.890 and their coefficients are biggest.So U_1 can be expressed mainly by these two variables. Similarly, V_1 can be expressed mainly by Y_7,Y_8 and Y_9.That is, we can study the correlations of the first pair canonical variates by studying the correlations among X_1,X_3,Y_7,Y_8 and Y_9. In the same way, the correlations of the second pair of canonical variates can be shown by the correlations among X_1,X_4,Y_1 and Y_8.

Since the first canonical coefficient is the biggest one, so the importance of pause (X_1) and self-repair (X_3) comes to the fore. Both pauses and self-revisions have greatest correlations with Y_7 (the number of phrases per T-unit), and the coefficient is -2.852; they have greater correlations with Y_8 (error free T-units to the total number of T-units), and the coefficient is 2.098; they have correlations with Y_9 (the number of clauses per T-unit), and the coefficient is -0.927.

The first pair of canonical variates shows the correlations among pause, self-repairs, grammatical accuracy and language complexity (lexical complexity and syntactic complexity). Self-repairs are correlated with language complexity positively and correlated with grammatical accuracy negatively. Pauses are correlated with language complexity negatively and are correlated with grammatical accuracy positively. These correlations reveal that self-repairs and pauses develop in opposite directions. But from the aspect of disfluencies traits, no matter the improvement of language complexity or grammatical accuracy, either of them will render the increase of disfluencies. This is in accordance with the findings of Bernstein Ratner (1977), which is if the language proficiency and the capabilities of grammar use develop in an unbalanced way, and students' attention will be transferred and will result in oral disfluencies.

Different language phenomena lead to different changes in disfluencies, either the increase or decrease of pauses or the decrease or increase of self-repairs. These correlations tell us that during these 2 years, students cannot control the language accuracy and language complexity at the same time. (1) When the number of phrases or chunks increases, the number of self-repairs rises as well. This rapid correction under psychological language control consciousness deduces the length of time for accurate and complex languages selection, so it deduces the production of pauses. We can infer from this tendency, with the improvement of language proficiency, the number of self-repairs will go down finally. Thus the use of chunks or phrases will decrease oral disfluencies of EFL students. It is similar to the findings of many researchers (Ping Yuan, 2010; Yan Chen, Qingqing, Zhao, 2010). (2) When the grammatical accuracy improves, the number of self-repairs will go down naturally for no necessity of revisions. The increase of this accuracy is gained by making the use of more pauses to earn more time for correct languages selection.

The correlations of the second pair of canonical variates primarily are expressed by the correlations among pauses(X_1, coefficient is -0.881), Y_8 (error free T-units to the total number of T-units, coefficient is -1.396) and Y_1 (simple sentences ratio, coefficient is -0.716). Pauses, simple sentences and grammatical accuracy are correlated positively. The more simple sentences used, the more pauses happened. In the same way, taking advantage of the use of pauses, grammatical accurate forms can be ensured, and finally the languages accuracy is improved.

Table 8. Canonical redundancy analysis

	Proportion of Variance of Set-1 Explained by its Own Can. Var.		Proportion of Variance of Set-1 Explained by its Opposite Can.Var.	
	Prop Var.	Prop accumulated	Prop Var.	Prop accumulated
CV1-1	.236	.236	.200	.200
CV1-2	.334	.570	.194	.394

	Proportion of Variance of Set-2 Explained by its Own Can. Var.		Proportion of Variance of Set-2 Explained by its Opposite Can.Var.	
	Prop Var.	Prop accumulated	Prop Var.	Prop accumulated
CV2-1	.236	.236	.201	.201
CV2-2	.133	.369	.077	.278

In canonical correlations analysis, we will discuss further about the proportion of variance explained by its opposite canonical variance. Because the variables we select have the biggest coefficients, so every variable can not only explain its own variance, but also explain variance in its opposite set. The higher coefficient, the more the variance information can be explained (Huixuan Gao, 2002). As shown in Table 8, the two variates of set 1 (disfluencies set) can explain accumulated 57% of own variance and accumulated 27.8% of their opposite variance. The two variates of set 2 (language acquisition set) can explain accumulated 36.9% of their own variance and accumulated 39.4% if their opposite variance. This proves that there are correlations between oral disfluencies traits of EFL students and their language acquisition development. But sole language acquisition development variables cannot explain all disfluencies traits and other factors, such as learning strategies, confidence, learning motivation and learning social environment, also have great effects on the development of disfluencies traits.

4. Discussion

From canonical correlation analyses above, we find the oral disfluencies traits of these 9 non-English major EFL college students have strong connections with their language acquisition development. The effects of comprehensive canonical correlation are better than the simple correlations among variables. In other words, all language acquisition variables working together affect disfluencies better. Through specific analyses, the obvious correlations are: pauses have close connections with simple sentences ratios, language complexity and grammatical accuracy. Self-repairs have tight associations with language complexity and grammatical accuracy. What is more, pauses and self-repairs develop in opposite ways. With the improvement of language proficiency, the language complexity goes up, and so do numbers of self-repairs. By the same token, grammatical accuracy increases, the number of pauses will increase as well. These kinds of correlations remind us that the increase of disfluencies is indispensable in the process of language acquisition. The increase of disfluencies is not necessarily the signal of the retreat of language proficiency. On the contrary, it can be regarded as a benign indicator of the improvement of language acquisition. With the improvement of language proficiency next round, the oral disfluencies will go down gradually, and oral language expressions will appear natural and smooth.

We also find that although the strong correlations between disfluencies traits and language acquisition development, parts of disfluencies traits cannot be explained. So if we want to study the comprehensive disfluencies traits of EFL students, we have to take other factors into account, such as learning motivations, learning strategies, self-confidence and social environment they live in and so on.

Acknowledgements

References

Baars, B., Motley, M., & MacKay, D. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382-391. http://dx.doi.org/10.1016/S0022-5371(75)80017-X

Bernstein Ratner, N. (1977). *Stuttering: A psycholinguistic perspective*. In R. F. Curlee, & G. M. Seigel (Eds.), *Nature and treatment of stuttering: New directions* (2nd ed., pp. 99-127). Boston: Allyn & Bacon.

Colburn，N., & Mysak, E. D. (1982). Developmental disfluency and emerging grammar II. Co-occurrence of disfluency with specified semantic-syntactic structures. *Journal of Speech and Hearing Research*, *25*, 421-427.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, *20*, 611-629. http://dx.doi.org/10.1016/S0022-5371(81)90202-4

Dollaghan, C. A., & Campbell, T. F. (1992). A procedure for classifying disruptions in spontaneous language samples. *Topics in Language Disorders*, *12*, 56-68. http://dx.doi.org/10.1097/00011363-199202000-00007

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*, 27-52. http://dx.doi.org/10.2307/412187

Galloway, V. (1987). From defining to developing proficiency. A new look at the decision in H. Byrnes (eds.): Defining and Developing Proficiency. *Guidelines, Implementations, and Concepts* (pp. 25-73). Lincolnwood, IL: National Textbook Company.

Garrett，M. F. (1975). The analysis of sentences production. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 133-177). New York, NY: Academic Press.

Gordon, P., Luper, H. L., & Peterson, H. A. (1986). The effect of syntactic complexity on the occurrence of disfluencies in 5-year-old non-stutterers. *Journal of Fluency Disorders*, *11*, 151-164. http://dx.doi.org/10.1016/0094-730X(86)90029-X

Higgs, T., & R. Clifford. (1982). The push towards communication. In T. V. Higgs (Ed.), *Curriculum, Competence, and the foreign Language Teacher* (pp. 57-79). Lincojnwood, IL: National Textbook Company.

Hotelling U. (1936). Relations between two sets of variables. *Biometrika, 28*, 321-377.

Huixuan Gao. (2002). Statistical analyses for multiple correlation variables of two sets (1). *Application of Statistics and Management, 21*(1), 57-64.

Hunt, K. W. (1970). Syntactic maturity in School children and adults. *Monographs of the Society for Research in Child Development* 35/1 (1, Serial No. 134). http://dx.doi.org/10.2307/1165818

Lee, L. L. (1974). *Developmental sentence analysis: A grammatical assessment procedure for speech and language clinicians*. Evanston, IL: Northwestern University Press.

Materson, J. J., & Kamhi, A. G. (1992). Linguistic trade-offs in school-age children with and without language disorders. *Journal of Speech and Hearing Research*, *35*, 1064-1075.

McNamara, T. (1996). *Measuring Second Language Performance*. London and New York: Longman.

Pearl, S. Z., & Bernthal, J. E. (1980). The effect of grammatical complexity upon disfluency behavior of non-stuttering preschool children. *Journal of Fluency Disorder*, *5,* 55-68. http://dx.doi.org/10.1016/0094-730X(80)90015-7

Ping Yuan, & Fenrong Guo. (2010). The correlations between chunks and Secondary language fluency. *Foreign Language World, 136*(1), 54-62.

Raupach, M. (1987). Temporal variables in first and second language speech production. In H. D. Decherd, & M. Raupach (Eds.), *Temporal Variables in Speech*. New York: Mouton.

Shirberg, E. (1994). Preliminaries to a Theory of Speech Disfluencies. Ph. D. thesis, University of California, Berkeley.

Starkweather, C. W. (1987). *Fluency and Stuttering*. Englewood Cliff, NJ: Prentice-Hall.

Tao Zhou. (2010). Teaching PPT: SPSS data statistical analysis and practices. Retrieved from http://www.ires.cn/Courses/SPSS

Wall, M. J., & Myers, F. L. (1984). *Clinical management of childhood stuttering*. Baltimore: University Park Press.

Yan Chen, & Qingqing Zhao. (2010). The correlations between chunks acquisition and oral communicative capabilities. *Shan Dong Foreign Language Teaching*, *135*(2), 52-55.