2012, Scienceline Publications
International Journal of Applied Linguistic Studies
Int. J. Appl. Ling. Stud., Vol. 1, No. 1: 01-05 (2012) ISSN 2322-5122

Analytic View toward Lexical Network Structure in Language Learning besides First Language

Ramin Sayedi* and Mehdi Azizi

Faculty of Maragheh Branch, Islamic Azad University (IAU), Maragheh, Iran

*E-mail: Ramin_Sayedi234@yahoo.com

Received: 05 Jun 2012; Accepted: 27 Aug 2012; Published: 15 Sep 2012

ORIGINAL ARTICLE
PII: S232251221200001-1

ABSTRACT:

This study is an attempt to demonstrate that an artificial neural network meant to simulate the potential learning mechanisms of second language learners was able to produce and not produce a small selection of nouns and verbs in a similar manner as L2 learners using four features related to lexical network models. This study helps support the potential for lexical network models to explain lexical production in L2 learners. Earlier studies have found empirical support for lexical networks in L2 learners. However, these past studies used either Boolean models or computational tools to investigate lexical growth. The studies did not use lexical features related to network models to simulate lexical production and learning. Thus, this study provides a broader perspective on how lexical features can inform lexical production. Lexical Network Theory asserts that the semantic portion of the lexicon is best seen as a network of word senses, where each sense is connected by links to other semantically-related senses of the same word, and, indirectly, to other words in the same semantic field. To this end, a neural network was trained to simulate L2 word production using a variety of word properties related to connectionist networks. Theories of connectionism and their links to artificial neural networks are relatively new. While neural network models exploring lexical acquisition in bilingual learners are common, few researchers in second language acquisition have examined neural network approaches to lexical production. When L2 neural network models have been explored, they have been in the absence of actual linguistic features or through the use of non-learning networks. Our purpose is to demonstrate how word properties that are linked to network models can be used to simulate word production by second language learners. We first did a corpus analysis of both L1 and L2 spoken discourse to select produced and unproduced words. Then we constructed an artificial neural network with the outputs for the words as either produced or unproduced and tested whether the network can correctly categorize the words based on word properties. The study demonstrates that artificial neural networks can categorize produced and unproduced words to a significant degree.

Key words: Hypernymy, Polysemy, Concreteness, Meaningfulness

INTRODUCTION

In many of past studies about lexical acquisition and lexical production we see a focus on broad measure of lexical growth such as lexical accuracy, lexical frequency, and lexical diversity (Polio, 2007). While these studies are very important they deal with surface level of linguistic features. On the other hand, connectionist perspective of lexical network gives us a broader understanding about the notion of lexical network. Lexical networks extend theories of lexical acquisition by giving a model of interconnections between words and not just memorizing words, their definition, orthography, and sound patterns. This theory claims that words interrelate with other words to form clusters of words which act categorically. These clusters connect to other clusters and they form the entire lexicons which are based on interconnections. These connections make it easy and possible for newly acquired words to be easily assimilated within these networks because words are not acquired in isolation. While learners progress lexically, they build lexical networks which are strengthened by differentiating sense relations between words and within words (Haastrup & Henriksen, 2000). Second language lexical production is very important because the inaccurate production of lexical items is the main factor in global errors that inhibit communication (Ellis, 1995) and lexical production is strongly related to academic achievement (Daller et al., 2003). A few recent studies have analyzed the development of L2 lexical networks (e.g. Crossley et al., 2008; Crossley et al., 1998), but there are very

To cite this paper: Sayedi R and Azizi M. 2012. Analytic View toward Lexical Network Structure in Language Learning besides First Language. *Int. J. Appl. Ling. Stud.* 1(1): 01-05.
Journal homepages: www.ijals.science-line.com

few studies like that. These studies show that L2 learners develop lexical networks over time, specifically in the development of hypernymic networks and word concreteness use (Crossley et al., in press), the development of semantic networks (Crossley et al., 2008), and polysemy knowledge (Schmitt, 1998).

This study shows that an artificial neural network meant to simulate the potential learning mechanisms of second language learners was able to produce and not produce a small selection of nouns and verbs in a similar manner as L2 learners using four features related to lexical network models. This study helps support the potential for lexical network models to explain lexical production in L2 learners. Earlier studies have found empirical support for lexical networks in L2 learners. However, these past studies used either Boolean models or computational tools to investigate lexical growth. The studies did not use lexical features related to network models to simulate lexical production and learning. Thus, this study provides a broader perspective on how lexical features can inform lexical production. Also, unlike past artificial neural network models in bilingual studies, this study examines which lexical features influence adult second language learning. While neural network models exploring lexical acquisition in bilingual learners are common, few researchers in second language acquisition have examined neural network approaches to lexical production. When L2 neural network models have been explored, they have been in the absence of actual linguistic features or through the use of non-learning networks (Meara, 2007). Our purpose is to demonstrate how word properties that are linked to network models can be used to simulate word production by second language learners. We first did a corpus analysis of both L1 and L2 spoken discourse to select produced and unproduced words. Then we constructed an artificial neural network with the outputs for the words as either produced or unproduced and tested whether the network can correctly categorize the words based on word properties. The study demonstrates that artificial neural networks can categorize produced and unproduced words to a significant degree.

METHODOLOGY

In this paper we try to study about produced and unproduced words in a way similar but simpler to that of beginning second language learners and we avoid all the complexities of an entire lexicon. To do this, we try to provide a simpler version of artificial neural network. Instead as suggested by Meara (2006), we intend to build a simple model that removes unnecessary complications under the premise that in emergent systems, simple connections lead to complex structures. Regardless of limitations in this study, it is functional in that it explores the properties of words inherent in the conceptual and the psycholinguistic properties inherent in the lexicon of the language users.

Data Collection

One of our requirements was to collect a list of frequent words produced by beginning L2 learners and another list of words that were not produced by second language learners but were produced by native speakers of English. We regarded the first list as words which were easier to produce than the second list which we mostly observed among native speakers. Then we needed two sorts of corpora (an L2 corpus and an L1 corpus). Also, we wanted to deal with natural language use then we aimed at both spoken and unprepared oral production. Our L2 corpus was collected in an English Language institution in Iran. It contained interviews of L2 learners who were taking an intensive course in English. Interview sessions were organized in a way to produce naturally accruing discourse. At first subjects were chosen from the lowest level of (level 1 & 2) of a 12 level program and then they were given a placement test. Also, to find out about unproduced words we needed a corpus of L1 speech of English native speakers. For this reason Santa Barb corpus (Du Bois et al., 2000) was selected. It consists of unprepared speech recordings of people in the United States in a natural setting. It was a rich and natural corpus which allowed us to have access to a big size of corpus (about 200,000 words).

In this study our major focus was on produced and unproduced words. For this reason we argue that words which are produced by our early learners are the ones which are easier to fully acquire. On the other hand, those words which were produced by native speakers and not by early L2 learners were the ones which are difficult to fully acquire. But we did not include all produced and unproduced words. We needed to consider four criteria in this study. The chosen words had to be produced by at least half of the L2 participants. The word also had to have a frequency above .10 in both the L1 and the L2 corpus. The word's use also had to fit clearly into a noun or verb categorization. If questions arose, the use of the word was analyzed in context to ensure its part of speech category. Finally, lexical values for the words in all examined categories (polysemy, hypernymy, concreteness, and meaningfulness) needed to be available. For this study, the first 10 verbs and nouns from each group that meet this criterion were selected. We collected a word list of 20 nouns including ten produced and ten unproduced nouns and 20 verbs including ten produced and ten unproduced verbs.

Word Measurement

In relation to produced and unproduced words we selected four word variables (polysemy, hypernymy, concreteness, meaningfulness). Polysemy and hypernymy are related to conceptual knowledge and concreteness and meaningfulness are related to psycholinguistic measures. And in one sense all these four variables are related to lexical networks.

Polysemous words are the words having two or more related meanings (e.g. foot, of a person, of bed, of maintain). In this manner, speakers will economize their vocabulary by extending words senses in order to conserve lexical storage space. In this way word meanings will be extended and words will possess multiple meanings. This is true for more frequent words (Zipf, 1945). In this view, words connect not only to a meaning, but also to networks of semantically similar words. In relation to polysemous words, lexical networks allow learners to identify meaning relationships between a word's senses (Verspoor & Lowie, 2003) because the word's senses are located within a single lexical item.

Hypernymy is regarded as a fundamental semantic relationship that is founded on the connection between general and specific lexical items (Haastrup & Henriksen, 2000). Hypernymic relations are hierarchical associations between hypernyms and hyponyms. This sort of relations allows for hierarchical categorizations which define how hyponyms inherit properties from their related hypernyms and allow set inclusion among category members. In this study, we determined hypernymy values using Word Net (Fellbaum, 1998) which is a lexical reference system inspired by current theories of lexical processing.

Concreteness refers to objects, materials, persons or any items that can be seen, heard, felt, smelled, or tasted and these words are more concrete than the others. Concrete words are advantageous for lexical acquisition because they are recalled more quickly and can be organized faster, and are comprehended more rapidly than abstract words (Paivio, 1991).

Meaningfulness refers to word associations. If a word is highly associated with other words, it is argued to be more meaningful. Associations such as meaningfulness are important for mediating the organization and memorization of words and afford for easier acquisition (Ellis & Beaton, 1993). In this study we determined concreteness and meaningfulness values using the Medical Research Council (MRC) psycholinguistic database (Wilson, 1988).

Analyses

Before we begin constructing and training our neural network, we conducted a t-test to see if there is a significant difference between the psycholinguistic and conceptual features of the words which we were to study them. Then we constructed a back propagation artificial neural network with four input nodes, two hidden nodes, and one output node; a bias node with a constant input of one was connected to the hidden and output nodes.

In this study we trained the entire dataset in the beginning because it was necessary to find the relevance of each lexical feature. Then we tested our artificial networks' accuracy on data it had never experienced. Since we had ten pairs of word we selected one pair in a way that one member in the pair be learned and the other one be unlearned verb. We removed this pair from our training set, and we assigned a random initial set of weights for each pair and then were trained on the other eighteen verbs, and were tested on the selected pair. To do this we needed ten separate runs of the program.

RESULTS

After performing a T-test we found a significant difference between Hypernymy values $t\ (1, 18)=-2.53$, $p< .05$, and concreteness value $t(1, 18)= -3.84$, $p< .001$. But there was no significance in verb groups in relation to other features. This test revealed a significant difference in word meaningfulness, $t(1, 18)= 2.58$, $p< 0.05$ with produced nouns showing higher meaningfulness values. But there was no significant difference in noun group in relation to other lexical features (see table 3).

Then a neural network was trained on the verb data set to enable the network to learn the correct classification. Results in each step were saved. Our trained net was to give us a classification for the 20 training data items. And all the verbs were classified correctly. Using NevProp we found a relevance of input for the classified verbs. Then a similar procedure was followed about the noun data set and again the net could classify all the nouns correctly. Again we found an input relevance for the classified nouns.

Table 1 - Selected verbs and their word features				
Verbs	**Polysemy**	**Hypernymy**	**Concreteness**	**Meaningfulness**
Eat	3	2	365	405
Do	38	1	359	347
Become	20	1	251	331
Grow	13	3	342	500
Want	35	1	337	430
Have	24	2	360	410
Get	34	1	402	558
Work	22	5	445	457
Hope	11	2	286	516
Come	37	5	464	376
Need	7	4	506	460
Read	5	4	419	508
Go	37	1	290	318
Enjoy	10	3	579	420
Like	14	2	371	371
Take	9	1	302	472
Call	11	5	556	389
Turn	21	3	502	430
Think	22	2	355	408
Buy	12	1	268	516

Table 2 - Selected nouns and their word features				
Noun	**Polysemy**	**Hypernymy**	**Concreteness**	**Meaningfulness**
Student	3	10	568	469
Man	17	7	595	533
House	6	5	540	612
Taxi	3	8	533	531
Year	8	18	472	513
Country	5	8	465	472
Food	19	7	365	531
Life	15	8	343	453
Engine	15	7	516	519
Mother	7	15	579	584
Dish	18	5	558	443
Chair	11	8	548	408
Team	9	15	594	554
office	4	11	582	608
Actor	9	7	332	393
Street	13	11	618	607
Friend	5	10	450	538
Star	1	11	586	490
Bath	8	8	339	337
Child	4	7	364	437

Table 3 - Means (Standard Deviations) for selected words			
Word Properties	**Variables**	**Produced**	**Un-produced**
Polysemy	Verbs	18.54 (11.80)	17.59 (11.98)
	Nouns	6.06 (4.04)	10.64 (6.52)
Hypernymy	Verbs	1.44 (1.05)	4.33 (1.56)
	Nouns	9.79 (3.32)	7.65 (1.70)
Concreteness	Verbs	335.00 (54.21)	456.11 (84.15)
	Nouns	507.79 (95.08)	476.90 (106.66)
Meaningfulness	Verbs	445.01 (75.02)	408.32 (37.01)
	Nouns	539.60 (66.09)	471.15 (65.81)

CONCLUSION

This study shows that an artificial neural network can simulate learning mechanisms of second language learners in a way to produce and not produce a small selection of nouns and verbs in a manner as L2 learners using four lexical features related to network models. Statistical analysis in this paper revealed that the produced and unproduced verbs differ in concreteness and hypernymy values and it proves the evidence that verb production might be influenced by these two features. Apart from that the noun groups differ in their meaningfulness values and this was supported by the input relevance found in the training set, and this proves the evidence that lexical features play an important role in L2 noun production. Besides our model's simplicity, this study represents which lexical features influenced adult language learning. We have made a distinction between learning verb and noun, while claiming that artificial neural network depends on different mechanisms for learning nouns and verbs. With reference to statistical analysis in this paper also we can support the notion that verbs which are less concrete are produced earlier, while nouns which are more meaningful and have more associations can be learned more easily. This is because concreteness is not an important aspect of verb production, and this means that more abstract words are learned first.

In this paper we see the importance of psycholinguistic features of lexicon. This study shows that psycholinguistic features may be indicative of whether a word is produced than conceptual features. Therefore, word production may not be a matter conceptual features which exist within the word, but it is based on psycholinguistic judgments of the characteristics of words.

This study is based on computational datasets which should allow for a scaling up toward natural languages. Also, this study is more than just describing words as nods in a net, but rather gives the nodes values which are taken from psycholinguistic and lexicographic datasets. It is true our model is simple but it is not limited by generalization. And the findings of this study provide evidence as to the relative strength of various lexical features I the production of L2 language by L2 learners. This can be considered as an important step in exploring which lexical features may influence word production in adult second language learners.

REFERENCES

Crossley, S. A., Salsbury, T. McCarthy, P. M. & McNamara, D. S. (2008). Using Latent Semantic Analysis to explore second language lexical development. In D. Wilson & G. Sutcliffe Eds., Proceedings of the 21st International Florida Artificial Intelligence Research Society pp. 136-141. Menlo Park, California: AAAI Press.

Crossley, S. A., Salsbury, T., & McNamara, D. S. (1998). in press. Measuring L2 lexical growth using hypernymic relationships. Language Learning.

Daller, H., van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. Applied Linguistics, 24 2, 197-222.

Du Bois, J. W., Chafe, W. L., Meyer, C. & Thompson, S. A. (2000). Santa Barbara Corpus of Spoken American English [CD-ROM]. Linguistic Data Consortium.

Ellis, R. (1995). Modified oral input and the acquisition of word meanings. Applied Linguistics, 16, 409-435.

Fellbaum, C. (1998). WordNet: An electronic lexical database. Cambridge, MA: MIT Press.

Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. International Journal of Applied Linguistics. 10 2, 221-240.

Meara, P. (2006). Emergent properties of multilingual lexicons. Applied Linguistics, 27 4, 620-644.

Meara, P. (2007). Growing a vocabulary. EUROSLA Yearbook, 7, 49–65.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45, 255–287.

Schmitt, N. (1998). Tracking the incremental acquisition of a second language vocabulary: A longitudinal study, Language Learning, 48 2, 281-317.

Polio, C. (2001). Research methodology in second language writing research: The case text-based studies. In T. Silva & P. K. Matsuda Eds. On second language writing. pp. 91-115. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Salsbury, T. (2000). The Acquisitional Grammaticalization of Unreal Conditionals and Modality in L2 English: A Longitudinal Perspective. Unpublished doctoral dissertation, Indiana University, Bloomington.

Verspoor, M., & Lowie, W. (2003). Making sense of polysemous words. Language Learning, 53 3, 547-586.

Wilson, M.D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. Behavioural Research Methods, Instruments and Computers, 201, 6-11.

Zipf, G. K. (1945). The meaning-frequency relationship of words. The Journal of General Psychology, 33, 251-256.