

# Differences of Early Semantic Relatedness between Late Talkers and Typically Developing Children

Yumi Kim<sup>a</sup>, Dongsun Yim<sup>b</sup>

<sup>a</sup>Language and Literacy Clinic Center Mindeulla, Seoul, Korea

<sup>b</sup>Department of Communication Disorders, Ewha Womans University, Seoul, Korea

**Correspondence:** Dongsun Yim, PhD  
Department of Communication Disorders, Ewha  
Womans University, 52 Ewhayecodae-gil,  
Seodamun-gu, Seoul 03760, Korea  
Tel: +82-2-3277-6720  
Fax: +82-2-3277-2122  
E-mail: sunyim@ewha.ac.kr

Received: October 4, 2018  
Revised: November 14, 2018  
Accepted: November 20, 2018

This work is based on a partial data from the first author's master thesis.

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (No. NRF-2016K2A9A2A19939734).

**Objectives:** Previous studies on late talkers (LT) suggested that residual linguistic deficits would remain in these children as they advance through the higher grades in school. This study investigated whether the residual linguistic deficits of LT exist from the infant and toddler stage by manifesting the differences of semantic relatedness between typically developing children (TD) and LT. **Methods:** Two hundred twenty-four reports of Korean MacArthur-Bates Communicative Development Inventories were used to investigate the semantic relatedness of infants and toddlers aged from 8 to 36 months. Semantic relatedness was measured by Jaccard's Index. After measuring the similarity of each group, one-way ANOVAs were conducted to investigate whether there were group differences. Three pairs of comparisons were made: infants' comprehension, toddlers' expression, and toddlers' comprehension. Infants' expression comparison was not conducted due to zero-expressive-words. **Results:** The differences between LT and TD infants were significant within the lexical categories in comprehension, except for 5 categories out of 18: 'Body parts', 'Households', 'Furniture and rooms', 'Places', and 'Quantifiers'. For toddlers, there was significant difference in every lexical semantic category both in expression and comprehension. All three pairs of comparison showed significant difference without categories. **Conclusion:** The current study results suggest that semantic relatedness is the key factor in word learning during infancy and that weak semantic relatedness in early age could lead to remaining linguistic residual deficits. Study results are discussed in regard to early intervention of 'at risk' children and suggest the direction of intervention for late talkers.

**Keywords:** Late talkers, Semantic relatedness, Language development, At-risk children, Early intervention, Semantic network

A language is a communication tool and a complex system for representing knowledge and thoughts. For successful communication, a speaker should select and produce appropriate words. When they select words to produce, a concept should be formed before the speech is made. Levelt (1989, 1991) explained the process of word speech as involving three steps: conceptualizer, formulator, and articulator. Since conceptualization and formulation should be preceded by a word phonetically produced, how the concepts are formed and related each other is critical for successful speech production.

The lexical retrieval process, which belongs to the formulator stage in Levelt's model, is also important because asymmetry exists between comprehension and production of early aged children. Early in development, children comprehend many more words than they produce. This is why those two tasks require different demands on retrieval processes (Gershkoff-Stowe & Hahn, 2007). Greater activation strengths are required to retrieve a word for producing than for comprehending a word (Capone & McGregor, 2005).

However, little is known about the lexical retrieval processes of

very young children. For comprehension, the listener responds to an auditory cue, and then the phonological representation previously stored in memory is activated. Finally, the activation spreads from the phonological level to the semantic level where the word is comprehended (Gershkoff-Stowe & Hahn, 2007). The retrieval of a word for production, however, is processed by the reverse flow. Its initial activation is derived from non-linguistic cues. Gershkoff-Stowe & Hahn (2007) explained that “these cues originates in semantic memory and spreads to the phonological level, and then to be accessed for production with sufficient strength” (p. 683). Either way, improvements in children’s ability to access stored representation in comprehension would support children’s comprehension ability which lead to generating the associated words.

There are several factors which influence this lexical retrieval access. Dell (1990) suggested that highly frequent words are processed faster. This can be explained by the fast mapping effect, the ability to learn and retain new words with only minimal exposure (Carey & Bartlett, 1978; Heibeck & Markman, 1987). Not only word frequency, but neighborhood density also affects the lexical retrieval access (Gershkoff-Stowe & Hahn, 2007).

Neighborhood density in lexical retrieval access can be divided into two sectors. One is phonological neighborhoods, and the other is semantic similarity. According to many studies of adults’ speech, lexical access is facilitated more by phonetically similar words than sparsely phonetically similar words (Harley & Bown, 1998). Levelt (1989) mentioned that words are also connected to other words at the semantic level. This semantic similarity is another potential in early word learning, when early-aged children acquire new words (Baldwin, 1992; Gershkoff-Stowe, Connell, & Smith, 2006; Smith & Yu, 2008).

Language develops continuously from the moment of being born, but each child has different language ability depending on their intrinsic capacities, environments, personalities, etc. Once an infant acquires a word, he starts to learn other words very fast, we call this phenomenon vocabulary burst (Bates et al., 1994). However, there are infants who do not show vocabulary burst or speak late. Late talkers have definite delays in language acquisition, as opposed to development in other areas (Robertson & Ellis Weismer, 1999).

Follow-up investigations of late talkers has proved the urgent necessity of intervention for late talkers. According to Robertson

and Ellis Weismer (1999), “the majority of children meet normative expectation on language assessment measures and measures of early reading skills by the elementary school-age period (Paul, 1996; Rescorla, Hadicke-Wiley, & Escarce, 1993), however they acquired significantly lower scores than control groups in various areas of linguistic functioning (Paul, 1996; Rescorla et al., 1993)” (p. 1235). Robertson and Ellis Weismer (1999) also supported the argument that “residual linguistic deficits” (p. 1235) would remain in these children as they advance through the higher grades in school.

This study thoroughly reviews the differences of semantic similarity between the typically developing children’s group (TD) and the late talkers group (LT). Studies have been conducted to measure the semantic relatedness among words, including feature similarity, and co-occurrence (Mirman & Magnuson, 2006). This study would also focus on the early semantic relatedness of infants and toddlers to verify the residual deficits of LT. The current research tries to verify the differences of words’ relatedness between TD and LT by counting co-occurrence data in Korean MacArthur-Bates Communicative Development Inventories (KM-BCDI; Pae & Kwak, 2011) based on Jaccard’s Similarity Index. Based on the semantic similarity, this study ultimately aims to compare semantic networks in each group and to investigate whether the features are same across two groups. Moreover, this study would analyze not only expressive vocabulary but also comprehensive words of children. Therefore, there would be 8 sectors: infants’ TD expression, infants’ TD comprehension, infants’ LT expression, infants’ LT comprehension, toddlers’ TD expression, toddlers’ TD comprehension, toddlers’ LT expression, and Toddlers’ LT comprehension.

First, the semantic relatedness of each word within lexical-semantic categories would be compared in each category. Secondly, the semantic relatedness of each word without lexical semantic categories would be examined. Finally, the visualization of the semantic network would be presented to compare two groups more easily.

---

## METHODS

### Participants

A total of 224 participants (55 infants and 169 toddlers) were as-

sessed using the K M-B CDI (Pae & Kwak, 2011). The K M-B CDI reports consists of two types: ‘words and gestures’ and ‘words and sentences’. Words and gestures are for infants from 8 to 17 months; words and sentences are for toddlers whose ages are from 18 months to 36 months.

Out of 224 reports, 196 reports were from the Wordbank database (<http://wordbank.stanford.edu/>). Thirty-eight participants were additionally recruited from daycare centers in Seoul, Seongnam, and Pusan in Korea. Among 224 participants, 31 were assessed as late talkers (7 were infants, and 24 were toddlers).

Participants were selected who meet the following criteria: (1) must have been born in Korea, (2) must use Korean at home with Korean speaking parents, and (3) must exhibit normal hearing and neurological development. The above information was collected through the Alberta Language Development Parent Questionnaire (ALDeQ; Paradis, Emmerzael, & Duncan, 2010). Table 1 shows the descriptive statistics of participants’ characteristics—number of participants, gender, age, and KM-BCDI score in expressive and comprehensive vocabularies.

The mean age for TD toddlers was 27 ± 6 months, for LT toddlers was 23 ± 4 months, for TD infants was 13 ± 2 months, for LT infants was 12 ± 2 months.

A set of one-way ANOVAs were conducted to investigate if there were group differences. There was no significant group difference in age in months, for both infants ( $F_{(1, 53)} = .452, p > .001$ ) and toddlers ( $F_{(1, 167)} = 7.260, p > .001$ ). There was no significant group difference in infants’ scores of KM-BCDI as well, for both expression ( $F_{(1, 53)} = 6.306, p > .001$ ) and comprehension ( $F_{(1, 53)} = .676, p > .001$ ). However, there was significant group difference in toddlers’ scores

of K M-B CDI, for both expression ( $F_{(1, 167)} = 64.798, p < .001$ ) and comprehension ( $F_{(1, 167)} = 26.488, p < .001$ ).

This comparison of average scores evidently shows the necessity of more sophisticated methods to assess infants and toddlers. In the period of infants and toddlers, they show cascading language development. The difference between words production may not seem serious in early age but it distinguishes LT from TD, and the difference remains.

### Measures

MacArthur–Bates Communicative Development Inventories

MacArthur-Bates Communicative Development Inventories (MBCDI; Dale & Fenson, 1996) is a widely-used assessment tool for assessing language development of infants and toddlers. Parents report which words their children produce on a checklist organized by lexical-semantic categories (Braginsky, Yurovsky, Marchman, & Frank, 2015). Pae & Kawk (2011) issued a Korean version of MBCDI, K M-B CDI.

Parental questionnaire: Alberta Language and Development Questionnaire

The ALDeQ (Paradis et al., 2010) was used to evaluate and check the overall developments of the participants based on the parental reports. This tool can be used for children who are not limited in language and cultural environment, and it is useful for screening children with language impairment (Restrepo, 1998). The questionnaire consisted of four parts: early milestone, current abilities in the first language, behavior patterns and activity preferences, and family history. Through these questionnaires, it was confirmed

**Table 1.** Participants' characteristics

	Toddlers		Infants	
	TD group (N = 145)	LT group (N = 24)	TD group (N = 48)	LT group (N = 7)
Sex				
Male	78	14	21	4
Female	67	10	27	3
Age (mo)	27 ± 6	23 ± 4	13 ± 2	12 ± 2
Average scores of KM-BCDI				
Expression	359.14 ± 195.13	36.29 ± 45.97	15.97 ± 16.70	0
Comprehension	464.02 ± 156.70	285.54 ± 161.43	99.5 ± 74.68	74.85 ± 68.88

Values are presented as mean ± SD.

TD = typically developing children; LT = late talker; KM-BCDI = Korean MacArthur-Bates Communicative Development Inventories (Pae & Kwak, 2011).

that participants met the criteria for this study.

**Data Scoring**

Parents were asked to check both comprehension and expression lists, and researchers double-checked those reports. If one expresses a word for example “kkokkio”, 1 was put, and if one does not express the word, 0 was put. The scores of comprehension and expression were separately inputted for each word. The lexical semantic categories were divided as K M-B CDI. One example of data arrangement is shown in Table 2.

After putting all answers to each word, the similarity was measured by the Jaccard similarity measure. When variables are binary, co-occurrences are measured by the Jaccard coefficient simi-

larity (Borg & Groenen, 2003). Jaccard’s Index is measured by the following formulation.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

In this study, Jaccard’s coefficient was measured through SPSS version 22.0 for Windows, both words within lexical-semantic category and words without categories were measured.

**Data Analysis**

The statistical analyses were conducted using SPSS version 22.0 for Windows as well. First, Jaccard’s coefficient was measured with absolute values. Three sets of one-way ANOVAs were conducted to compare group scores for similarity.

**Table 2.** Data coding for LT toddlers sounds category

	꺅꺅꺅 Quack quack	꺅꺅꺅 Oink	꺅꺅꺅 Knock knock	꺅꺅꺅 Woof woof	꺅꺅꺅 Beef beef	꺅꺅꺅 Meow	꺅꺅꺅 Yum yum	꺅꺅꺅 Eoheung	꺅꺅꺅 Moo	꺅꺅꺅 Twitting	꺅꺅꺅 Choo choo
TO4001	0	1	1	1	1	1	1	1	1	0	0
TO4002	1	0	1	1	0	1	0	1	0	0	0
TO4003	1	0	0	1	0	0	1	1	0	0	0
TO4004	1	1	1	1	0	1	0	0	0	1	0
TO4005	0	0	1	0	0	0	1	0	0	0	0
TO4006	0	0	0	1	0	0	0	1	0	0	0
TO4007	0	0	0	1	0	1	0	1	0	0	0
TO4008	0	0	0	0	0	0	0	0	0	0	0
TO4009	0	0	0	0	0	0	0	0	0	0	0
TO4010	0	0	0	0	0	0	0	0	0	0	0
TO4011	1	0	1	1	1	1	1	1	1	0	1
TO4012	1	1	1	1	1	1	0	1	1	1	1

LT=late talker.

**Table 3.** Similarity matrix of ‘sounds’ category for TD toddlers’ expression

	꺅꺅꺅 Quack quack	꺅꺅꺅 Oink	꺅꺅꺅 Knock knock	꺅꺅꺅 Woof woof	꺅꺅꺅 Beef beef	꺅꺅꺅 Meow	꺅꺅꺅 Yum yum	꺅꺅꺅 Eoheung	꺅꺅꺅 Moo	꺅꺅꺅 Twitting	꺅꺅꺅 Choo choo
꺅꺅꺅 Quack quack	1.00	.935	.865	.917	.901	.915	.858	.908	.883	.884	.884
꺅꺅꺅 Oink	.935	1.00	.886	.924	.921	.895	.865	.915	.850	.878	.864
꺅꺅꺅 Knock knock	.865	.886	1.00	.895	.879	.866	.836	.886	.847	.848	.835
꺅꺅꺅 Woof woof	.917	.924	.895	1.00	.930	.958	.901	.924	.860	.874	.861
꺅꺅꺅 Beef beef	.901	.921	.879	.930	1.00	.915	.871	.908	.897	.857	.912
꺅꺅꺅 Meow	.915	.895	.866	.958	.915	1.00	.872	.895	.844	.885	.858
꺅꺅꺅 Yum yum	.858	.865	.836	.901	.871	.872	1.00	.865	.839	.854	.827
꺅꺅꺅 (a growl sound)	.908	.915	.886	.924	.908	.895	.865	1.00	.891	.905	.864
꺅꺅꺅 Moo	.883	.850	.847	.860	.897	.844	.839	.891	1.00	.866	.923
꺅꺅꺅 Twitting	.884	.878	.848	.874	.857	.885	.854	.905	.866	1.000	.867
꺅꺅꺅 Choo choo	.884	.864	.835	.861	.912	.858	.827	.864	.923	.867	1.000

TD=typically developing children.

## RESULTS

### Group Differences in Similarity within Lexical–Semantic Categories

One hundred sixty-eight matrices were abstracted. The K M-B CDI for toddlers (words and sentences) has 24 semantic categories, therefore, 24 matrices for TD toddlers' expression and comprehension, respectively, and for LT toddlers as well. The K M-B CDI for infants (words and gestures) has 18 categories, therefore, 18 matrices for TD infants' expression and comprehension, respectively, and for LT infants as well. Since LT infants do not express even a single word, the comparison on similarities of infants' expression was not possible. Therefore the comparison of infants' expression was not conducted.

One example of similarity matrix is shown in Table 3.

The descriptive statistics of the average similarity of infants' comprehension are provided in Table 4.

To compare the difference of similarity of infants' comprehension between TD and LT on each lexical semantic category, a one-way ANOVA was conducted with 'group' as the within-subjects factor.

**Table 4.** Mean similarity of infants' comprehension

Category	TD group	LT group
1. Sounds (14)	.548±.172	.391±.344
2. Vehicles (7)	.454±.275	.232±.375
3. Animals (21)	.561±.153	.431±.392
4. Body parts (20)	.435±.231	.431±.371
5. Clothing (10)	.498±.247	.603±.316
6. Toys (10)	.452±.233	.286±.376
7. Food drink (35)	.357±.209	.176±.296
8. Households (16)	.415±.203	.327±.348
9. Furniture rooms (9)	.578±.182	.502±.284
10. Places (6)	.468±.268	.389±.406
11. Outside (12)	.549±.223	.373±.451
12. People (17)	.338±.222	.251±.332
13. Routines (19)	.655±.171	.702±.182
14. Pronouns (7)	.595±.197	.382±.388
15. Quantifiers (6)	.496±.237	.527±.387
16. Action words (43)	.477±.173	.397±.316
17. Descriptive words (20)	.446±.185	.350±.356
18. Function words (12)	.551±.174	.778±.373

Values are presented as mean ± SD and the number of infants in parenthesis. TD = typically developing children; LT = late talker.

Fifteen categories out of eighteen were significantly different ( $p < .001$ ) between those two groups listed above. In 5 categories, which included 'Body parts', 'Households', 'Furniture and rooms', 'Places', and 'Quantifiers', the differences were not significant.

The group differences of similarity of infants' comprehension between TD and LT group in each lexical category are shown in Figure 1.

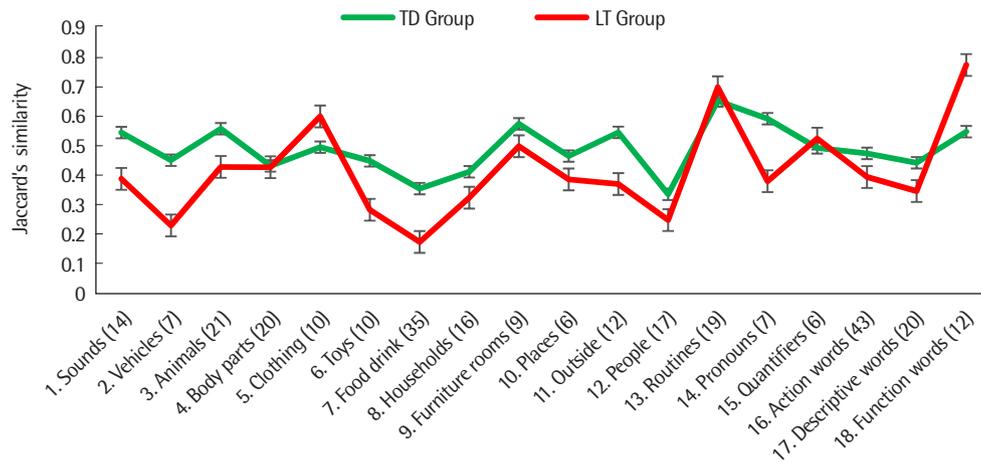
The descriptive statistics of the mean similarity of toddlers' expression are provided in Table 5.

To compare the difference of similarity of the toddlers' expression between the TD and LT group on each lexical semantic category, a one-way ANOVA was conducted with 'group' as the within-subjects factor. Twenty-three categories out of 24 were significantly different ( $p < .001$ ) between those two groups listed above. The mean difference of similarity in 'Helping verbs' category does not exist, for the expression scores were all 0 for LT group. The group

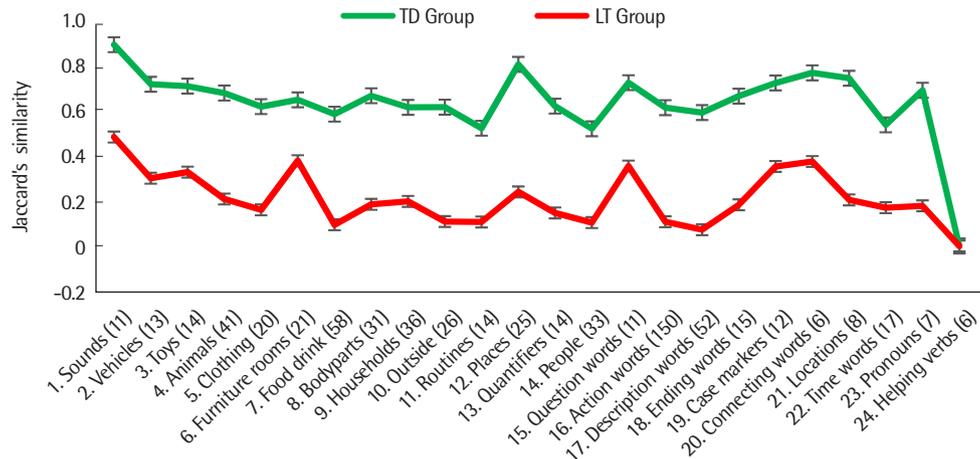
**Table 5.** Mean similarity of toddlers' expression

Category	TD group	LT group
1. Sounds (11)	.893±.043	.483±.231
2. Vehicles (13)	.717±.146	.301±.343
3. Toys (14)	.709±.120	.328±.311
4. Animals (41)	.679±.173	.208±.269
5. Clothing (20)	.618±.199	.161±.287
6. Furniture rooms (21)	.648±.217	.378±.395
7. Food drink (58)	.585±.187	.094±.224
8. Body parts (31)	.667±.164	.185±.293
9. Households (36)	.615±.188	.198±.352
10. Outside (26)	.616±.182	.108±.272
11. Routines (14)	.522±.231	.106±.278
12. Places (25)	.806±.098	.241±.304
13. Quantifiers (14)	.621±.174	.146±.314
14. People (33)	.520±.181	.103±.237
15. Question words (11)	.724±.231	.355±.486
16. Action words (150)	.614±.158	.107±.276
17. Descriptive words (52)	.592±.134	.072±.238
18. Ending words (15)	.665±.165	.183±.388
19. Case markers (12)	.723±.209	.353±.485
20. Connecting words (6)	.726±.195	.167±.000
21. Locations (8)	.744±.136	.204±.393
22. Time words (17)	.537±.178	.169±.341
23. Pronouns (7)	.691±.147	.178±.373
24. Helping verbs (6)	0	0

Values are presented as mean ± SD and the number of toddlers in parenthesis. TD = typically developing children; LT = late talker.



**Figure 1.** Mean similarity of infants' comprehension. TD=typically developing children; LT=late talker.



**Figure 2.** Mean similarity of toddlers' expression. TD=typically developing children; LT=late talker.

differences of similarity of toddlers' expression between TD and LT group in each lexical category are shown in Figure 2.

The descriptive statistics of the average similarity of toddlers' comprehension are provided in Table 7.

To compare the difference of similarity of toddlers' comprehension between the TD and LT group on each lexical semantic category, a one-way ANOVA was conducted with 'group' as the within-subjects factor. Twenty-four categories out of 24 were significantly different ( $p < .001$ ) between those two groups listed above. The group differences of similarity of toddlers' comprehension between TD and LT group in each lexical category are shown in Figure 3 and Table 6.

### Group Differences in Similarity without Lexical-Semantic Categories in KM-B CDI

The similarities between words without lexical categories were measured. Six hundred and forty-one words for toddlers' expression and comprehension, and 284 words for infants' expression and comprehension, were conducted. Since the average score of LT infants' expression is 0, the comparison of the similarity of infants' expression was not possible.

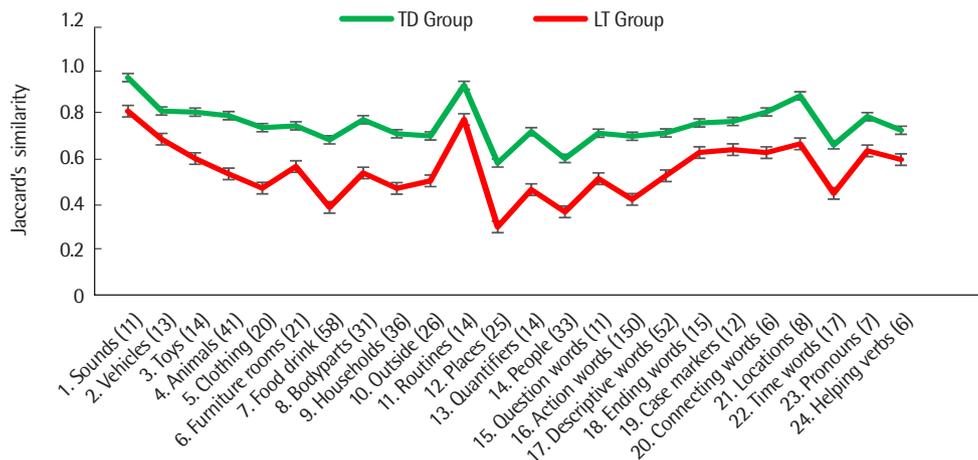
The descriptive statistics of the average similarity of infants' comprehension without lexical semantic categories are provided in Table 7.

To compare the difference of similarity of infants' comprehension between TD group and LT group without lexical semantic category, a one-way ANOVA was conducted with 'group' as the

**Table 6.** Mean similarity of toddlers' comprehension

Category	TD group	LT group
1. Sounds (11)	.952 ± .023	.805 ± .105
2. Vehicles (13)	.806 ± .115	.683 ± .156
3. Toys (14)	.801 ± .098	.598 ± .236
4. Animals (41)	.786 ± .136	.531 ± .201
5. Clothing (20)	.733 ± .152	.468 ± .243
6. Furniture rooms (21)	.741 ± .172	.563 ± .257
7. Food drink (58)	.680 ± .160	.384 ± .207
8. Body parts (31)	.768 ± .130	.534 ± .241
9. Households (36)	.706 ± .160	.467 ± .212
10. Outside (26)	.697 ± .151	.501 ± .198
11. Routines (14)	.918 ± .056	.769 ± .124
12. Places (25)	.578 ± .170	.298 ± .249
13. Quantifiers (14)	.715 ± .121	.462 ± .242
14. People (33)	.598 ± .158	.364 ± .202
15. Question words (11)	.709 ± .119	.510 ± .225
16. Action words (150)	.695 ± .128	.419 ± .186
17. Descriptive words (52)	.710 ± .098	.523 ± .157
18. Ending words (15)	.754 ± .084	.625 ± .206
19. Case markers (12)	.761 ± .092	.637 ± .206
20. Connecting words (6)	.802 ± .098	.624 ± .232
21. Locations (8)	.872 ± .059	.662 ± .186
22. Time words (17)	.658 ± .134	.445 ± .244
23. Pronouns (7)	.781 ± .103	.632 ± .185
24. Helping verbs (6)	.722 ± .141	.594 ± .233

Values are presented as mean ± SD and the number of toddlers in parenthesis. TD = typically developing children; LT = late talker.



**Figure 3.** Mean similarity of toddlers' comprehension. TD = typically developing children; LT = late talker.

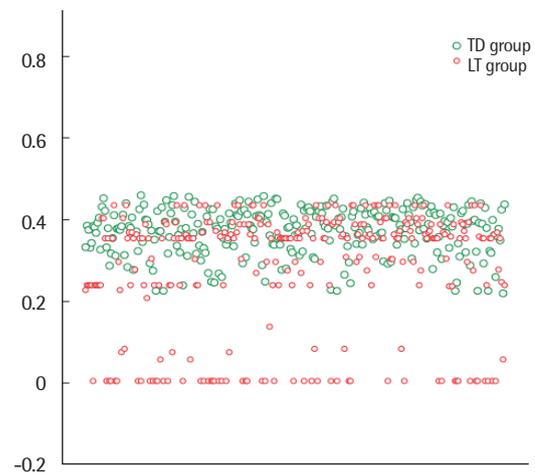
within-subjects factors.

Without lexical semantic categories, two groups were significantly different ( $p < .001$ ) as listed above. The group differences of similarity of infants' comprehension between TD and LT group without the lexical semantic category are shown in Figure 4.

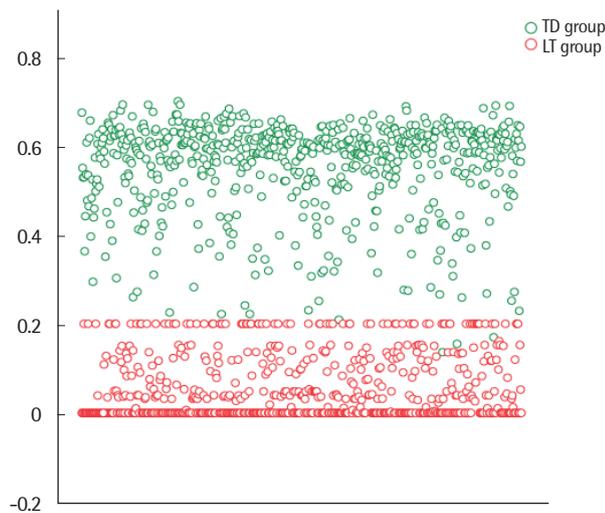
**Table 7.** Mean similarity of expression and comprehension without lexical-semantic categories

	TD group	LT group
Infants comprehension	.372 ± .153	.283 ± .307
Toddlers expression	.564 ± .160	.077 ± .202
Toddlers comprehension	.668 ± .138	.399 ± .187

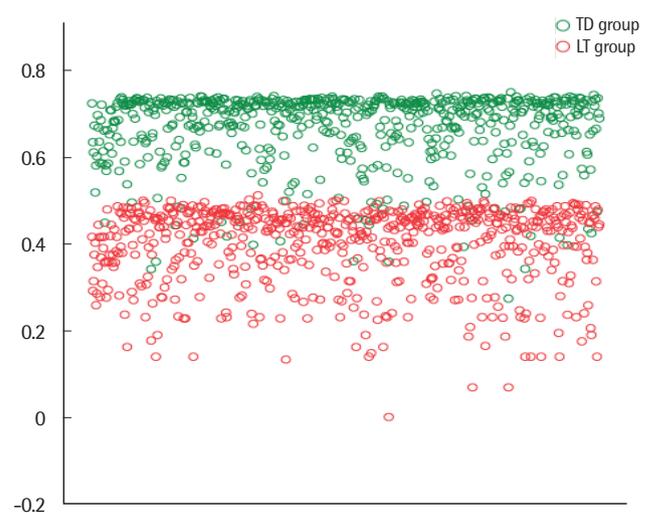
Values are presented as mean ± SD. TD = typically developing children; LT = late talker.



**Figure 4.** Mean similarity of infants' comprehension without categories. TD = typically developing children; LT = late talker.



**Figure 5.** Mean similarity of toddlers' expression without categories. TD = typically developing children; LT = late talker.



**Figure 6.** Mean similarity of toddlers' comprehension without categories. TD = typically developing children; LT = late talker.

The descriptive statistics of the average similarity of toddlers' expression without lexical semantic categories are provided in Table 7.

Without lexical semantic categories, the two groups were significantly different ( $p < .001$ ) as listed above. The group differences of similarity of toddlers' expression between the TD and LT group without lexical semantic categories are shown in Figure 5.

The descriptive statistics of the average similarity of toddlers' comprehension without lexical semantic categories are provided in Table 7.

To compare the difference of similarity of toddlers' comprehension between the TD and LT group without lexical semantic category, a one-way ANOVA was conducted with 'group' as the within-subjects factor. Without lexical semantic categories, the two groups were significantly different ( $p < .001$ ) as listed above. The group differences of similarity of toddlers' comprehension between TD and LT group without lexical semantic categories are shown in Figure 6.

## CONCLUSION

Late talkers are assessed to have obvious delays in acquiring language, in contrast to development in other areas (Robertson & Ellis Weismer, 1999). Approximately half of the late talkers catch up with their peers by age three (Paul, 1991; Rescorla & Schwartz, 1990; Rescorla, Roberts, & Dahlsgaard, 1997). Even though they

reach the normal range of linguistic ability during the elementary school-aged period, they receive significantly lower scores in language assessment than their peers in the higher grades in school (Paul, 1996; Rescorla et al., 1993). Paul, Hernandez, Taylor, & Johnson (1996) insisted that residual linguistic deficits would remain in these children as they advance through the higher grades in school.

This study hypothesized the residual deficits start from conceptualization stage. Out of theories on the representation of knowledge, this study is based on the Spreading activation model. The model is explained as a complex network of associations. Specific items are distributed in the conceptual space with related concepts, and the concepts are linked by associations (Solso, MacLin, & MacLin, 2008).

To examine the linkage of semantics, the current study compared the similarity of words between LT and TD. A total of 224 reports of K M-B CDI (Pae & Kwak, 2011) were used to investigate the semantic relatedness of infants and toddlers from 8 to 36 months of age. One hundred ninety-six reports were from Wordbank (<http://wordbank.stanford.edu/>), and 38 reports were gathered by additional recruitment. K M-B CDI has two versions: one is for infants (8 to 17 months) and the other is for toddlers (18 to 36 months). Out of 224, 55 were infants, and 169 were toddlers. Among them, 7 infants were LT, and 24 toddlers were LT.

Semantic relatedness was measured by Jaccard's Index. After measuring the similarity of each group, one-way ANOVAs were

conducted to investigate whether there were group differences. Three pairs of comparisons were made: infants' comprehension, toddlers' expression, and toddlers' comprehension. Infants' expression comparison was not conducted due to zero-expressive words.

#### Differences of Semantic Relatedness within Categories

K M-B CDI (Pae & Kwak, 2011) has its lexical semantic categories. Words and gestures for infants have 18 lexical semantic categories, and words and sentences for toddlers has 24.

First, the mean similarity of infants' comprehension was compared. The mean similarity of words in fifteen categories was significantly different. TD's mean similarity was higher than LT's. This indicates that relatedness of LT's is significantly weaker than TD's. The 5 categories which did not show significance were 'Body parts', 'Households', 'Furniture and rooms', 'Places', and 'Quantifiers'.

Early vocabularies tend to consist of nouns while verbs and closed-class forms are typically acquired later (Bates et al., 1994). According to Hills, Maouene, Maouene, Sheya, & Smith (2009), the fact that "early noun learning is slow at first and becomes fast" suggests that "already learned nouns might help in learning new nouns" (p. 729). This means that both infants and toddlers should show their increased numbers of acquired words, once they learned how to speak a few words. However, the difference of semantic relatedness has become more significant as infants get older. Even though the comparison of a few categories did not show significant differences in infants' comprehension, that all the categories of toddlers showed significant differences could give further evidence of residual deficits in late talkers.

#### Differences of Semantic Relatedness without Categories

All three pairs resulted in having significant differences in comparison without categories. The most outstanding difference was evident in toddlers' expressive vocabulary, and the most obscure difference was observed in infants' comprehension.

However, there were variables which were laid at the bottom in the infants' comprehension graph. These variables were not distinctively evident in the within categories comparison. These variables, laid at the bottom, could be the evidence of residual deficits

which lead late talkers to linguistic delay. Those lowest similarity words ( $p < .004$ ) in LT infants are arranged in Appendix 1.

#### Clinical Implications and Limitations

Approximately half of the late talkers catch up with their peers by age three, and these children are designated 'late bloomers' (Paul, 1991; Rescorla & Schwartz, 1990; Rescorla et al., 1997). Even after vocabulary reaches the normal range, a number of late talkers continue to demonstrate production delays in other areas, including phonology, morphology, syntax and narrative abilities (Paul, 1991; Paul et al., 1996; Paul & Smith, 1993; Rescorla & Schwartz, 1990; Robertz, Rescorla, Giroux, & Stevens, 1998). Furthermore, follow-up study on late talkers has revealed that even though they reach the normal range of linguistic ability in the elementary school-aged period, they receive significant lower scores in language assessment than their peers in the higher grades in school. This suggests that there could be linguistic residual deficits in late talkers. The present study hypothesized that there are linguistic residual deficits in late talkers and tried to verify the evidence of residual deficits by comparing the similarity of TD and LT.

The idea of comparing the similarity of two groups is based on the Spreading Activation Model. The Spreading Activation Model explains semantic processing with the network model. According to the Spreading Activation Model, concepts which are related to each other activate stronger and faster. Collins & Loftus (1975) created this model, and they suggested that concepts are featured by nodes in a semantic network. Concepts which have a similar meaning are connected with each other within the network.

Early links between language and cognition provide the foundation for processing information (Ferguson & Waxman, 2017). A study of Smith & Yu (2008) provided that the link between language and categories is established early in infancy. Categorization is fundamental to word learning, which makes it easier to learn novel words (Chomsky, 2011).

The hypothesis of this study was that the residual deficits of LT came from the weak ability of semantic relatedness. In the semantic networks, links between words imply relatedness. Categorizing ability, which connects concepts towards languages, could be the reason for the linguistic residual deficits of LT. Both TD and LT showed lower similarity index in the comparison of without cate-

gories than within categories. The words laid in the bottom in the similarity graph (Figure 4) can also be the evidence of the weak categorizing ability of infants.

This study did not offer the learning mechanism of words via semantic networks. The small number of LT infants was a weakness in this study. Due to the small number of LT infant subjects, the results in some categories were distorted; such as ‘Function words’, which showed higher similarity in LT than TD. Another limitation of this study is that the words inventory should not have been limited to the K M-B CDI. There could have been more words that LT could speak, but those words were not considered.

However, the current study proposed that semantic relatedness is a key factor in word learning in infancy and that weak semantic relatedness in early age could lead to remaining linguistic residual deficits. Therefore, this study can be the academic evidence for urging early intervention of ‘at risk’ children and suggests the direction of intervention on late talkers.

## REFERENCES

- Baldwin, D. A. (1992). Clarifying the role of shape in children's taxonomic assumption. *Journal of Experimental Child Psychology*, 54, 392-416.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., ... & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, 21, 85-123.
- Borg, I., & Groenen, P. (2005). Modern multidimensional scaling: theory and applications. *Journal of Educational Measurement*, 40, 277-280.
- Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2015). Developmental changes in the relationship between grammar and the lexicon. *Proceedings of the Cognitive Science Society*, 256-261.
- Capone, N. C., & McGregor, K. K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research*, 48, 1468-1480.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference*, 15, 17-29.
- Chomsky, N. (2011). Language and other cognitive systems. What is special about language? *Language Learning and Development*, 7, 263-278.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125-127.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313-349.
- Ferguson, B., & Waxman, S. (2017). Linking language and categorization in infancy. *Journal of Child Language*, 44, 527-552.
- Gershkoff-Stowe, L., Connell, B., & Smith, L. (2006). Priming overgeneralizations in two- and four-year-old children. *Journal of Child Language*, 33, 461-486.
- Gershkoff-Stowe, L., & Hahn, E. R. (2007). Fast mapping skills in the developing lexicon. *Journal of Speech, Language, and Hearing Research*, 50, 682-697.
- Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology*, 89, 151-174.
- Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: an examination of fast mapping. *Child Development*, 58, 1021-1034.
- Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? *Psychological Science*, 20, 729-739.
- Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. (1991). *Lexical access in speech production*. Cambridge, MA: Blackwell.
- Mirman, D., & Magnuson, J. S. (2006). The impact of semantic neighborhood density on semantic access. In *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 1823-1828). Mahwah, NJ: Erlbaum.
- Pae, S., & Kwak, K. C. (2011). *Korean MacArthur-Bates Communicative Development Inventories (K M-B CDI)*. Seoul: Mindpress.
- Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: using parent report on first language development. *Journal of Communication Disorders*, 43, 474-497.
- Paul, R. (1991). Profiles of toddlers with slow expressive language development. *Topics in Language Disorders*, 11, 1-14.
- Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology*, 5, 5-21.
- Paul, R., Hernandez, R., Taylor, L., & Johnson, K. (1996). Narrative development in late talkers: early school age. *Journal of Speech, Language, and Hearing Research*, 39, 1295-1303.
- Paul, R., & Smith, R. L. (1993). Narrative skills in 4-year-olds with normal,

- impaired, and late-developing language. *Journal of Speech, Language, and Hearing Research*, 36, 592-598.
- Rescorla, L., Hadicke-Wiley, M., & Escarce, E. (1993). Epidemiological investigation of expressive language delay at age two. *First Language*, 13, 5-22.
- Rescorla, L., Roberts, J., & Dahlsgaard, K. (1997). Late talkers at 2: outcome at age 3. *Journal of Speech, Language, and Hearing Research*, 40, 556-566.
- Rescorla, L., & Schwartz, E. (1990). Outcome of toddlers with specific expressive language delay. *Applied Psycholinguistics*, 11, 393-407.
- Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, 41, 1398-1411.
- Roberts, J., Rescorla, L., Giroux, J., & Stevens, L. (1998). Phonological skills of children with specific expressive language impairment (SLI-E). *Journal of Speech, Language, and Hearing Research*, 41, 374-384.
- Robertson, S. B., & Ellis Weismer, S. (1999). Effects of treatment on linguistic and social skills in toddlers with delayed language development. *Journal of Speech, Language, and Hearing Research*, 42, 1234-1248.
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106, 1558-1568.
- Solso, R. L., MacLin, M. K., & MacLin, O. H. (2008). *Cognitive psychology*. Boston, MA: Allyn and Bacon.

**Appendix 1.** The lowest mean similarity words of late talker infants

Words	Category	Words	Category
꼬끼오(cock-a-doodle-doo)	Sounds	배(Pair)	Food drink
윙(Bzzz)	Sounds	쌀(Rice)	Food drink
배(Boat)	vehicles	오렌지(Orange)	Food drink
비행기(Plane)	Vehicles	주스(Juice)	Food drink
택시(Taxi)	Vehicles	초콜렛(Chocolate)	Food drink
벌(Bee)	Animals	돈(Money)	Households
쥐(Mouse)	Animals	가위(Scissor)	Households
하마(Hippo)	Animals	포도(Grapes)	Food drink
목(Neck)	Body parts	콩(Beans)	Food drink
똥꼬(Ass)	Body parts	칼(Knife)	Food drink
등(Back)	Body parts	학교(School)	Place
피(Blood)	Body parts	땅(Land)	Outside
반지(Ring)	Toys	시소(Seesaw)	Outside
사진(Picture)	Toys	누나(Older sister)	People
감(Perssimon)	Food drink	선생님(Teacher)	People
총(Gun)	Toys	아저씨(Mr.)	People
굴(Oysters)	Food drink	아줌마(Mrs.)	People
국(Soup)	Food drink	나(내) (I, me)	Pronouns
꿀(Honey)	Food drink	거기(There)	Pronouns
껌(Gum)	Food drink	깎아(Cut)	Action words
김치(Kimchi)	Food drink	(계단)내려가(Step down)	Action words
라면(Instant noodle)	Food drink	때려(Hit)	Action words
떡(Rice cake)	Food drink	빼(Subtract)	Action words
무우(Radish)	Food drink	많아(Many)	Descriptive words
밤(먹는 것) (Chestnut)	Food drink	더워(Warm)	Descriptive words
같아(똑같아) (Same)	Descriptive words	무서워(Scary)	Descriptive words
~말구(without)	Function words	추워(Cold)	Descriptive words

## 국문초록

### 말늦은 아동과 일반아동 간의 의미 연결성의 차이

김유미<sup>1</sup>(언어재활사, 제1저자) · 임동선<sup>2</sup>(교수, 교신저자)

<sup>1</sup>언어인지연구소 민들레, <sup>2</sup>이화여자대학교 일반대학원 언어병리학과

**배경 및 목적:** 말늦은 아동에 대한 후속 연구는 말늦은 아동에게 잔존하는 언어적 결손이 있다는 것을 제시하였다. 본 연구는 말늦은 아동의 이러한 언어적 결손이 어휘 개념의 저장된 형태에서 온 것이라고 보고, 말늦은 아동 그룹과 일반아동 그룹 간의 의미 연결성의 차이를 설명하고자 한다. **방법:** 영아 55명 중 7명, 유아 169명 중 24명이 말늦은 아동으로 평가되었다. 말늦은 아동 그룹과 일반아동 그룹 간의 표현어휘와 이해어휘의 유사도의 차이를 일원배치분산분석(one-way ANOVA)을 통해 비교하였다. **결과:** 범주별로 비교하였을 때, 영아 이해어휘는 18개의 범주 중 13개의 범주에서 유의미한 차이를 보였다. 유아의 경우 표현어휘와 이해어휘 모두 일반아동과 말늦은 아동 간 유의미한 차이를 보였다. 범주를 제거한 단어의 유사성 비교에서도 세 비교군 모두 유의미한 차이를 보였다. **논의 및 결론:** 본 연구에서 말늦은 아동의 잔존하는 언어적 결함이 의미적 연결성에서 보여진다는 것을 확인하였으며, 복잡한 의미적 연결성이 어휘의 이해와 표현 모두에서 정의 상관관계를 보인다는 것을 알 수 있다. 이를 통해 고위험군(at risk)에 있는 영유아의 언어증재의 중요성을 시사하는 바이다.

**핵심어:** 말늦은 아동, 의미 연결성, 언어 발달, 고위험군 아동, 조기 중재, 의미 연결망

본 논문은 제1저자의 석사학위논문의 일부를 발췌한 것임.

본 논문은 2017년도 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. NRF-2016K2A9A2A19939734).