DOES TALKER VARIABILITY HELP ADULTS LEARN NOVEL WORDS?

Sandy Abu El Adas[1], Ivy Yen[2], Susannah V. Levi[2]

[1]Basque Center on Cognition, Brain and Language, [2]The Department of Communicative Sciences and Disorders, New York University
[1]s.abueladas@bcbl.eu

ABSTRACT

Exposure to multiple talkers has been shown to be beneficial for adults learning speech sounds in a second language and for children learning novel words in a lab-based word learning task. The current study examines the extent to which variability is beneficial for adults in a word learning task and how individual differences in reading and language skills affect performance. Seventy-six participants were randomly assigned to either the multiple-talker (four talkers) or single-talker condition and learned novel object-nonword pairings. To assess word learning, listeners completed three naming tasks throughout the study (after 4 exposures, after 8 exposures, and after a brief delay) in which they had to generate the name of the object. Overall, listeners with higher language skills showed better word learning than those with poorer language skills. No effects of reading ability or talker variability were found. These results suggest that variability may only be beneficial in some tasks.

Keywords: speech processing, talker processing, phonetic variability, individual differences.

1. INTRODUCTION

In the domain of spoken language processing, different areas of research have examined the benefit of variability on learning. The most widely studied area is learning non-native sound contrasts. Numerous studies have shown that when adults learn a novel non-native sound contrast, exposure to multiple talkers improves both perception and production [1–5]. For example, Japanese speakers trained on the English /r/-/l/ contrast in the multiple talker condition performed better than those in the single talker condition, demonstrating the importance of variability on the development of phonemic categories [5]. The explanation for this benefit is that talker variability allows listeners to determine which variability is tied to the phonemic contrast and which is not [6], and that this type of perceptual warping helps reduce the focus on irrelevant acoustic information [5].

Not all studies, however, found a benefit of multiple talkers for learning non-native sound contrasts (e.g., [7]). The lack of consistency across studies may be explained by differences in the stimuli used, whether listeners are tested on new items (generalization), amount of exposure during training, and age of the participants (infants, children, adults).

While many studies have examined the benefits of variability on L2 sound learning, fewer studies have examined its effects on word learning. Richtsmeier et al. [8] investigated the effect of talker variability (single talker vs. 10 talkers) on word learning in pre-school children and found that the multiple talker condition resulted in more accurate and faster naming. A similar study of children with and without Developmental Language Disorder (DLD) also found a benefit of talker variability [9]. Together, these two studies suggest that the exposure to multiple talkers facilitates word learning in children, presumably by helping the learner create abstract phonological representations of the novel words.

In addition to talker variability, phonological factors such as phonological complexity and phonotactic probability have been shown to impact word learning [10–11]. For example, studies show that adults perform better on a word learning task when words are shorter and less phonologically complex [10]. Phonotactic probability also plays a role on word learning. However, there are inconsistencies across children and adults. While children show a high phonotactic probability advantage [11], adults show a high phonotactic probability disadvantage [12].

Individual factors such as language and reading ability also impact word learning. Studies show that individuals with poor reading or language skills perform differently [13–15]. However, studies have examined these skills separately. It is important to examine both reading and language skills simultaneously (within the same individuals) because there is an overlap in the underlying skill differences in phonological processing that then impact word learning.

While most research on the effects of talker variability in word learning focused on children that have not yet reached phonological maturity, little is known on how adults cope with talker variability when learning new words and whether this varies based on individual differences in the domains of reading and language ability. This is particularly important in the context of studies showing inconsistent results of how people process and use talker variability and how age plays a role on whether variability facilitates learning [16].

In the current study, we address this gap by examining how talker variability and individual differences in reading and language skills affect word learning in adults. In addition, we investigate how phonological information (phonotactic probability) affects word learning given the literature showing a relationship between phonological processing skills and word learning.

2. METHODS

2.1. Participants

Seventy-six participants ages 18-40 (M=30.8, SD=6.47) completed the study. All listeners were native speakers of U.S. English with no history of speech, language, or hearing disorders. All participants passed a web-based headphone check [17]. Participants were recruited through Prolific (www.prolific.co) and paid $20.

2.2. Stimuli

Sixteen CVC nonwords were selected from Storkel and Hoover's database [18]. Half had high phonotactic probability (M=0.0096, SD=0.0031, range=0.0075-0.0172) and half had low phonotactic probability (M=0.0032, SD=0.0001, range=0.0019-0.0048). Phonotactic probability was determined by the biphone frequency corresponding to the adult corpus. Each nonword was paired with an object from the Novel Object and Unusual Name (NOUN) Database [19].

Four female native speakers of U.S. English ages 19-24 from the New York City area produced the nonwords. Recordings took place in a sound-attenuated booth with a SHURE 10A head-mounted microphone and a Fostex FR-2LE recorder using a 44.1 kHZ sampling rate. All items were amplitude normalized for the word learning task.

2.2. Procedure

Participants were randomly assigned to one of two conditions: (a) high talker variability (n=29) or (b) low talker variability training (n=47). Those in the high variability condition heard the nonwords produced by all four talkers, and those in the low variability training were randomly assigned to one of the four talkers. Participants also completed additional tasks described below. The experiment took approximately one hour to complete and was hosted on Gorilla experiment builder [20].

2.2.1. Word learning task

During the training phase, participants heard the nonword-object pair, repeated the nonword aloud, and then answered a yes/no question about the semantic features of the object (e.g., "Is the object black and white?") to check engagement (Figure 1).

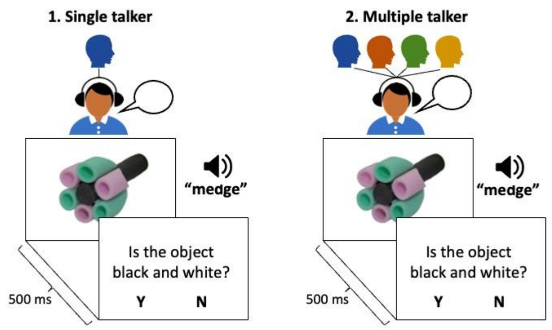

Figure 1: Procedure for **(1)** single-talker and **(2)** multiple-talker conditions in the word learning task.

During *Training 1* participants heard the 16 nonwords repeated 4 times in random order (16*4=64 trials), followed by a naming task (*Naming 1*), where they were presented with the object and asked to say the name out loud. They then completed a second identical training block (*Training 2*) and naming task (*Naming 2*). Following a brief delay (approximately 10 minutes) in which they completed filler tasks, participants then completed a final naming task (*Naming 3*).

2.2.2. Individual difference measures

Language ability was assessed using a modified version of the Recalling Sentences subtest of the CELF-4 [21]. Participants heard pre-recorded sentences and were instructed to repeat them. All scoring procedures from the CELF-4 were followed. Reading ability was assessed with the Rapid Online Assessment of Reading (ROAR), a quick and automatic online assessment previously tested with both adults and children [22-23]. In this task, participants complete a lexical decision task with orthographic real words and nonwords.

One limitation is that the score distribution on the ROAR was positively skewed with more participants with higher scores (see Table 1). This is expected given that our participants were typically developing adults.

Assessment	Mean (SD)	Range
Recalling Sentences	84.17 (10.91)	44-96
ROAR	78.76 (4.26)	58-84

Table 1: Mean, standard deviation (SD), and range for the language and reading measures.

3. RESULTS

A logistic mixed-effect model was fit to the naming accuracy data (coded as 0,1) using the lme4 package [24] in RStudio [25-26]. The fixed effects included naming (1, 2, 3), condition (multiple talkers, single talker), language ability, reading ability, phonotactic probability (high, low), a 3-way interaction between naming, condition, and phonotactic probability along with their 2-way interactions, and also 2-way interactions for condition by reading ability and condition by language ability. Raw scores (scaled and centered) were used for our individual measures (Recalling Sentences, ROAR) to allow for more variability in the data. The model also included random intercepts for participant and nonword (item). All categorical fixed effects (phonotactic probability, task, condition) were sum-coded.

The analysis revealed a significant effect of naming and the emmeans analysis [27] showed that participants performed significantly better on Naming 2 and 3 compared to Naming 1, and that there was a significant decline in performance between Naming 2 and Naming 3 (Table 2 and Figure 2).

Contrast	*Estimate*	*SE*	*z-ratio*	*p**
1 vs. 2	-2.17	0.119	-18.17	**<.001**
1 vs. 3	-1.93	0.118	-16.29	**<.001**
2 vs. 3	0.24	0.099	2.43	**0.040**

Table 2: emmeans analysis of naming. *p* values are adjusted using Holm's method and indicated as *p**

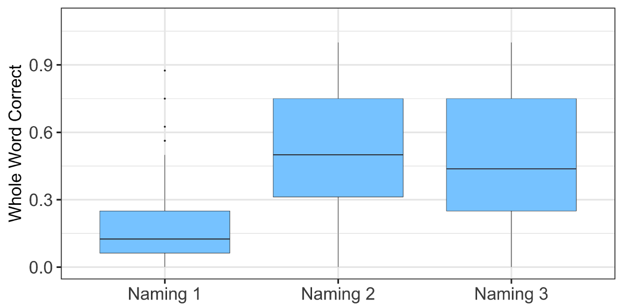

Figure 2: Naming by proportion correct.

In addition, there was a significant effect of language ability (β=0.55, *SE*=0.16, *p*<0.001), with better naming for participants with higher language scores (see Figure 3).

In addition, the analysis revealed a significant interaction between condition and phonotactic probability (β=-0.10, *SE*=0.04, *p*=0.0239). No pairwise comparisons from the emmeans analysis for condition within each of the levels of phonotactic probability reached significance, but the interaction suggests that the difference for condition was greater in the high than the low phonotactic probability items (see Figure 4).

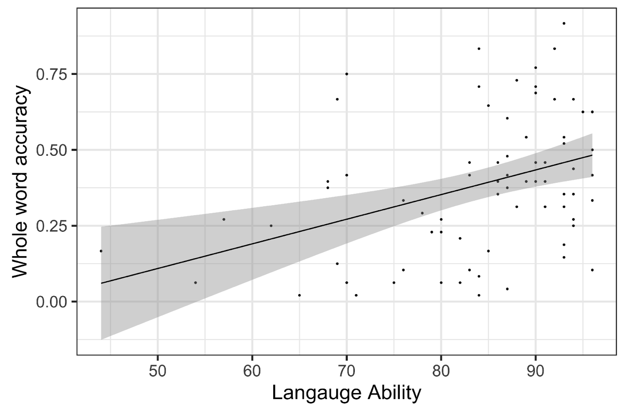

Figure 3: Relationship between language ability (raw scores on Recalling Sentences) and whole nonword accuracy.

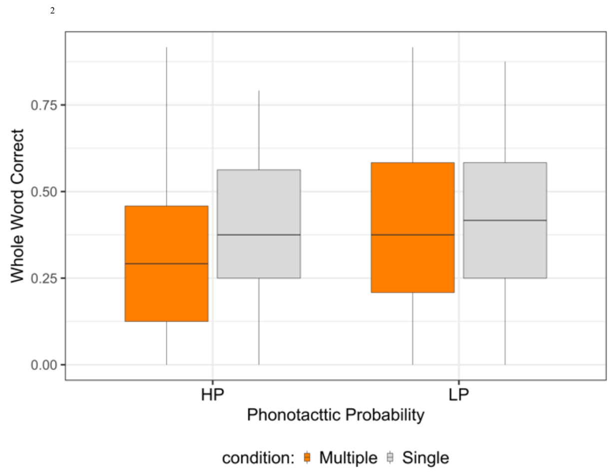

Figure 4: Proportion correct by condition (multiple=orange, single=grey) and phonotactic probability (HP=high, LP=low).

Unexpectedly, there was no effect of talker variability (β=0.05, *SE*=0.16, *p*=0.75) or reading ability (β=0.24, *SE*=0.17, *p*=0.15). None of the other effects or interactions reached significance.

4. DISCUSSION

In the current study, we examined the extent to which talker variability modulates performance on a word learning task with adults, and whether individual differences in reading or language ability influence performance. Consistent with previous studies, our results revealed an effect of exposure on word learning, where participants showed gains in learning over time and performed better after two exposure blocks compared to after only one. However, these gains were not fully maintained after a brief delay. This decrease might be explained by the fact that the participants were in the early stages of learning, where integration of the newly learned words have not been fully integrated in the lexicon.

Consistent with previous studies on word learning (e.g., [14]), we found that participants with higher language skills performed better on the naming tasks. Of note is that this study with children examined learning of low frequency words rather than nonwords. Here, we show that language ability plays an important role on word learning even after carefully controlling the stimuli.

Contrary to our hypothesis and previous studies (e.g., [13]), we found no effect of reading ability on word learning. There are several possible explanations for the difference between our study and previous studies. First, the type of task we used to assess reading ability may have been less sensitive to reveal differences in reading ability among adults who – as far as we know – had typical reading development. Here, we used the ROAR, which required participants to read the items and decide if the items are real words, not a decoding task. Second, it is possible that the lack of reading ability effect on word learning is related to statistical power and the overall variance of the ROAR scores. In fact, most of our participants scored high on the ROAR, which may have reduced the effects of reading on word learning.

Also surprisingly, talker variability did not facilitate learning, and participants in the single- and multiple-talker conditions performed similarly on the word learning task. These results go against the word learning studies with children that found a benefit for talker variability. One possible reason for this discrepancy may relate to the number of talkers used during training. The two studies with children used ten talkers in the multiple talker condition, while the current study used only four. Apart from the quantity of talkers, the variability of the talkers themselves may also play a role. While both our study and the previous studies used only female speakers, it may be possible that talker variability gains depend on how much acoustic differences exist between talkers. Indeed, in the current study, all female speakers were from the same age range and had a similar dialect, reducing the amount of acoustic variability between speakers. More research is needed to understand the effects of quantity and quality of talker variability on word learning.

Another possible reason for the lack of benefit of talker variability may relate to the amount of exposure during training. It is possible that our listeners did not have enough exposure to the talkers to show a talker variability benefit. This idea is in line with the effect of naming block on word learning where participants did not maintain their learning after a short delay. This is also consistent with a previous study showing that at the early stages of learning, variability can slow processing and reduce the differences between multiple and single talker conditions, but after sufficient exposure, variability is beneficial [28].

Finally, it is possible we did not find any talker variability effects due to age. The two previous studies that found a talker effect on word learning [8-9] tested children, whereas our study tested adults. Studies show age influence on talker processing where listeners become more sensitive to talkers and talker differences in their native language compared to talkers in an unfamiliar language [e.g., 29].

Most of the research on talker variability in word learning has focused on children that have not yet reached an adult-like phonological maturity. Here, we add to a body of literature on talker variability in word learning by testing adults on a set of phonologically legal nonwords. Because our stimuli contained nonwords that are all phototactically legal, it is possible that we did not find differences between single- and multiple-talker conditions because our participants have matured phonological capabilities that washed-out the differences between the groups. The interaction between condition and phonotactic probability suggests that individuals show talker variability benefits depending on how phonologically familiar nonwords are. While post-hoc testing did not find any significant differences between the two groups, there was a greater difference between the multiple and single talker conditions on nonword with high phonotactic probability. This may suggest that when the newly learned words are more word-like, talker variability negatively impacts learning, especially in the early stages of learning. These results are consistent with previous studies on word learning in adults showing better performances on items from low phonotactic probability compared to high phonotactic probability [12].

5. CONCLUSION

Our study found no benefit of talker variability in adults in a word learning paradigm, although this might have been driven by differences in design and stimuli. In addition, language ability, but not reading ability, influenced learning. Participants with higher language ability performed better on the naming tasks (regardless of training condition). We also found that phonological familiarity impacts whether talker variability is helpful during a word learning task, with better gains in words that are more phonologically dissimilar to other words in the language.

6. REFERENCES

[1] Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura Y.I. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299–2310.

[2] Kartushina, K., & Martin, C.D. (2019). Talker and Acoustic Variability in Learning to Produce Nonnative Sounds: Evidence from Articulatory Training. *Language Learning*, *69*(1), 71–105.

[3] Lambacher, S. G., Martens, W. L., Kakehi, K. , Marasinghe, C. A. , & Molholt, G.. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics,* 26(*2*), 227–247.

[4] Lively, S.E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English/r/and/l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America,* 96(4), 2076–2087.

[5] Logan, J. S., Lively, S. E. & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: a first report. *The Journal of the Acoustical Society of America,* 89(*2*), 874–886.

[6] Holt, L. L. & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America,* 119(5), 3059–3071.

[7] Brosseau-Lapré, F., Rvachew, S., Clayards, M. & Dickson, D. (2013). Stimulus variability and perceptual learning of nonnative vowel categories. *Applied Psycholinguistics*, 34(*3*), 419–441.

[8] Richtsmeier, P. T., Gerken, L., Goffman, L. & Hogan, T.P. (2009). Statistical frequency in perception affects children's lexical production. *Cognition*, 111(3), pp. 372–377.

[9] Plante, E., Bahl, M. , Vance, R. & Gerken, L. (2011). Beyond phonotactic frequency: Presentation frequency effects word productions in specific language impairment. *Journal of Communication Disorders,* 44(*1*), 91–102.

[10] Papagno, C. & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *The Quarterly Journal of Experimental Psychology Section A*, 44(*1*), 47–67.

[11] Storkel, H. L. & Rogers, M. A. (2000). The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics & Phonetics*, 14(*6*), 407–425.

[12] Storkel, H. L., Armbrüster, J. & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech Language and Hearing Research*.

[13] Alt, M., Hogan, T., Green, S., Gray, S., Cabbage, K. & Cowan, N. (2017). Word learning deficits in children with dyslexia. *Journal of Speech Language Hearing and Research*, 60(*4*),1012–1028.

[14] Windfuhr, K. L. & Snowling, M. J. (2001). The Relationship between Paired Associate Learning and Phonological Skills in Normally Developing Readers. *Journal of Experimental Child Psychology*, 80(2), 160–173.

[15] McGregor,K., Gordon, K., Eden, N., Arbisi-Kelm, T. & Oleson, J. (2017). Encoding deficits impede word learning and memory in adults with developmental language disorders. *Journal of Speech Language and Hearing Research*, 60(*10*), 2891–2905.

[16] Quam, C., & Creel, S. C. (2021). Impacts of acoustic-phonetic variability on perceptual development for spoken language: A review. *Wiley Interdisciplincary Reviews: Cognitive Sciences*, 12(*5*), e1558.

[17] Woods, K. J., Siegel, M. H., Traer, J. & McDermott J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention Perception & Psychophysics.*, 79(*7*), 2064–2072.

[18] Storkel, H. L. & Hoover, J. R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavioural Research. Methods*, 42(*2*), 497–506.

[19] Horst, J. S. & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavioural Research Methods*, 48(*4*), 1393–1409.

[20] Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N. & Evershed J. (2018). Gorillas in our midst: Gorilla. sc. *Behavioural Research Methods*.

[21] Semel, E., Wiig, E. H. &. Secord, W. A. (2003). Clinical evaluation of language fundamentals (CELF-4). *San Antonio TX*: Psychological Corporation.

[22] Yeatman J. D., et al. (2021). Rapid online assessment of reading ability. *Scientific Reports*, 11(*1*),1–11.

[23] Amrita, B., Annaliese, B., Tau, N., Pablo, R. & Saloni, K. (2020). The role of intrinsic reward in adolescent word learning.

[24] Bates, D., Maechler, K., Bolker, B. & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014*.

[25] Team, R. C. (2017). *R: A language and environment for statistical computing.*

[26] RStudio Team. (2016). *RStudio Team. RStudio: Integrated Development for R.*

[27] Lenth, R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means (Version R Package Version 1.3.4).

[28] Lee D. & Baese-Berk, M. (2021). Non-native English listeners' adaptation to native English speakers. *JASA Express Letters.*, 1(*10*), 105201.

[29] Levi, S. V., & Schwartz, R. G. (2013). The Development of Language-Specific and Language-Independent Talker Processing. *Journal of Speech, Language, and Hearing Research*, *56*(3), 913–925.