The Effect of Social Network Size on Hashtag Adoption on Twitter

Iris Monster,[a] Shiri Lev-Ari[b,c]

[a]*Behavioural Science Institute, Radboud University*
[b]*Department of Psychology, Royal Holloway University of London*
[c]*Department of Psychology of Language, Max Planck Institute for Psycholinguistics*

Abstract

Propagation of novel linguistic terms is an important aspect of language use and language change. Here, we test how social network size influences people's likelihood of adopting novel labels by examining hashtag use on Twitter. Specifically, we test whether following fewer Twitter users leads to more varied and malleable hashtag use on Twitter, because each followed user is ascribed greater weight and thus exerts greater influence on the following user. Focusing on Dutch users tweeting about the terrorist attack in Brussels in 2016, we show that people who follow fewer other users use a larger number of unique hashtags to refer to the event, reflecting greater malleability and variability in use. These results have implications for theories of language learning, language use, and language change.

Keywords: Twitter; Social network size; Hashtags; Language use; Language change

1. Introduction

Imagine that your favorite bar has changed its furniture and installed seats that look in between chairs and bar stools. If you only heard a couple of your friends refer to these seats, and one of them, Tess, used the term *chairs*, you might be inclined to use that term for them as well. In contrast, if you heard a dozen of your friends refer to these pieces of furniture, it is likely that you will be less influenced by Tess's specific lexical choice, as you will consider the lexical choices of all your friends. In other words, the weight that you give the input from each speaker is inversely related to the number of speakers you receive input from. The more sources you have, the less informative each source is, and

Correspondence should be sent to Shiri Lev-Ari, Department of Psychology, Royal Holloway University of London, Egham TW20 0EX, UK. E-mail: shiri.levarei@rhul.ac.uk

the less you will be influenced by each source. One interesting potential consequence of that is that people who interact with fewer others should be more influenced by each person they interact with. Consequently, people who interact with fewer others should have more malleable linguistic representations and more malleable language use. This is the hypothesis that this paper tests by examining hashtag use on Twitter. This hypothesis also has implications for language change as it suggests that people who interact with fewer other people might play an important role in the propagation of linguistic change.

1.1. Learning statistics from the environment

People are sensitive to the frequency and distribution of input in their environment. For example, high-frequency words are easier to access, as reflected in faster recognition time (e.g., Forster, 1976; Luce, 1986) and shorter production latencies (Forster & Chambers, 1973). They are also easier to remember, enabling more high-frequency words than low-frequency words to be kept in short-term memory (Roodenrys, Hulme, Lethbridge, Hinton, & Nimmo, 2002). Similarly, the relative frequency of a word's competitors influences its processing and production. Thus, words take longer to identify if they have a phonological or orthographic neighbor of higher frequency (e.g., Grainger, O'regan, Jacobs, & Segui, 1989). Likewise, the more frequent meaning of homonyms is accessed more quickly than their less frequent meaning, at least in the absence of any context (Simpson & Burgess, 1985), and production of a subordinate label is slower than of its dominant alternative (Lev-Ari & Shao, 2017). These frequency effects fit with an exemplar model account of language processing (Goldinger, 1998; Johnson, 1997) that proposes that we store each token that we hear, and later activate it upon hearing a related token. Activation in this model is therefore proportional to the number of tokens that we stored. Bayesian models of language processing and learning also posit that our processing and representations are sensitive to the frequency in our input by modeling our likelihood of producing or interpreting a word or a phoneme correctly as dependent on its conditional probability in the context (e.g., Kleinschmidt & Jaeger, 2015). Frequency, then, plays an important role in language learning and processing, and influences our linguistic representation. This suggests that it is likely to play a role in language change as well. Indeed, models of language change often likewise assume that speakers are more likely to produce the most frequent token in their input.

Importantly, there is evidence that suggests that we are sensitive not only to the frequency of a certain token or behavior, but also to the number of individuals exhibiting that behavior. This sensitivity starts at an early age (Corriveau, Fusaro, & Harris, 2009; Herrmann, Legare, Harris, & Whitehouse, 2013; Schillaci & Kelemen, 2013) and extends even to some non-human animals (Haun, Rekers, & Tomasello, 2012). Thus, both children and chimpanzees, though not orangutans, are more likely to copy a behavior that is produced by each of three conspecifics once than a behavior produced by one conspecific three times (Haun et al., 2012). Furthermore, it is not only the raw number of individuals that exhibit each behavior that is crucial but the relative number of individuals exhibiting each of the competing behaviors. Thus, when referring to the hybrid bar furniture, if upon

hearing two speakers use the term *chair* you would be more likely to copy it if you have only heard one speaker use the competing term *bar stool* than if you have heard fifty speakers use the competing term *bar stool*. Interestingly, this suggests that people who communicate with fewer other people should be more influenced by each speaker they hear because that speaker would constitute a higher proportion of their social network. People who interact with fewer others would thus require input from fewer speakers before their representations and use would exhibit noticeable change.

Previous research provides some evidence for the greater malleability of people with smaller social networks (Lev-Ari, 2017, 2018). In particular, in one study (Lev-Ari, 2017) participants reported how many people they talked to in a typical week, and then they performed a modified version of a perceptual learning task (Norris, McQueen, & Cutler, 2003). In this task participants were exposed to atypical pronunciations which had a Voice Onset Time that was ambiguous between /d/ and /t/. In one condition, the context disambiguated the phonemes as /d/, and in the other as /t/. Later on, participants were tested on their boundary between /d/ and /t/ with a new speaker, in order to measure the degree to which they changed their general representations. Results showed that the smaller participants' social network, the more they showed perceptual learning, that is, a shift in their boundary in line with their exposure. To ensure that the effect of social network size is not due to differences in how motivated participants were to do the task, or in how good they are in learning a speaker's speech patterns, half of the participants were tested with the same speaker rather than with a new speaker. In that case, social network size should not influence the magnitude of the perceptual learning effect, since everyone, regardless of social network size, had the same experience with the speaker. Indeed, in this case, everyone showed perceptual learning and social network size did not modulate the magnitude of this effect. A follow-up study (Lev-Ari, 2018) pursued the same question using a similar paradigm, but at the lexical level. In this case, participants' boundary between *some* and *many* was tested. As in the phonological study, participants who reported to speak to fewer people per week shifted their general boundary between *some* and *many* more. Also, as in the phonological study, social network size did not influence ability to learn the exposure speaker's use of *some* and *many*, indicating that the effect of social network size is not due to differences in motivation to do the task or in ability to learn the speech patterns, but in likelihood of generalizing the learning. Additionally, in this latter experiment, participants' production was also tested. As with perception, the smaller participants' social network, the more they shifted the boundary between *some* and *many* in their own productions. These studies thus indicate that social network size influences how susceptible people are to the influence of a new speaker.

1.2. *Propagation of linguistic change*

The relative malleability of speakers' linguistic representations and use has implications for language change. Synchronic variation is frequent in language at all linguistic levels: Different speakers might produce the same phoneme differently, they might alternate between comparable syntactic structures across situations, and they might refer to

the same object differently, for example, naming it *sofa* while others say *couch*. Language change occurs when new variants enter the language or when an existing variant increases in frequency or spreads across the language community beyond its original group of users or contexts of use. One of the puzzles of such spread is the *Threshold Problem*, how a rare variant propagates across the linguistic community if people tend to repeat frequent variants (Nettle, 1999). One proposed account is that prestige can lead to the propagation of rare variants despite their rarity (Nettle, 1999). Other research suggests that a change in the population, such as an increase in L2 speakers, can influence which variants are preferred as well (Dale & Lupyan, 2012; Hills & Adelman, 2015). Here, we propose that another way by which rare variants can propagate is via people with a small social network. While it might seem that people with smaller social networks might have lower social capital and would thus be less likely to influence others, previous research shows that the spread of information as well as trends often depends on weak ties, which are by definition network members whose social centrality in one's network is small (Granovetter, 1973). We will return to this point in the discussion.

2. Study

The goal of this study was to further test whether people with smaller social networks are more malleable in their linguistic use. Unlike the previous experiments, this study examines natural and unmanipulated linguistic exposure and language use, and thus tests the ecological validity of the previous findings. It also focuses on production, whereas previous studies focused mainly or only on perception. Additionally, this study examines language use on Twitter, and thus provides a more reliable and objective measure of social network, as the network of Twitter users is available publicly.

One feature of Twitter that is particularly useful for our purposes is the use of hashtags. Hashtags are strings preceded by a # that classify the content and relevance of the tweet. Common uses of hashtags include reference to an event that the tweet discusses, such as *#Rio2016* or *#FrenchElections*, pointing to the person or the topic the tweet refers to, as in *#Bowie* or *#ObamaCare*, or expression of an opinion, as in *#LoveTrumpsHate*. Some of these common uses of hashtags render them similar to referring expressions, and similarly to such expressions, there are often several ones that compete. For example, *#Olympics2016* refers to the same event as *#Rio2016*. We decided to focus on hashtag use in our study, because, unlike common lexical items, hashtags are not used in real-world interactions. Therefore, monitoring the Twitter input that a Twitter user receives provides us with their exposure to the competing terms. In this study we focused on the hashtags surrounding one specific event, the terrorist attacks in Brussels on March 22, 2016. By focusing on a specific unanticipated event, we can be sure that we are tracking the full exposure to hashtags about the event, as Twitter users could have not seen any relevant hashtags prior to March 22, 2016. To test whether users with smaller social networks are more susceptible to the influence of others' language use, we collected all exposure and use information for a subset of Dutch speakers during the 2 weeks

following the attacks. We measured how many different hashtags these users used over the 2-week period. We predicted that Twitter users who follow *fewer* other users would show *more* variability in hashtag use overall.

2.1. Data collection

In order to test our hypothesis, we first needed to determine which hashtags were used to describe the attacks. To do so, we used Twitter's Streaming API to collect and examine all English tweets tweeted during the day following the attacks. We lemmatized them using the English Stanford parser (Klein & Manning, 2003; Manning et al., 2014). We then ran two Term Frequency–Inverse Document Frequency algorithms. These algorithms test the likelihood of each hashtag that occurred to appear in relevant versus irrelevant tweets, and thus can generate a list of relevant hashtags. To ensure best performance, we ran two different algorithms, based on two different definitions of relevance—time-based and topic-based. The time-based definition of relevance defined any tweet that was tweeted in the 24 h prior to the attacks (i.e., 8am CET on March 21, 2016–8am CET on March 22, 2016) as irrelevant, and any tweet tweeted during the 24 following the first attack (i.e., 8am CET on March 22, 2016–8am CET on March 23, 2016) as relevant. The topic-based definition of relevance defined any tweet with the word *Brussels* tweeted within the 24 h following the attack as relevant, and any tweet in that time period without this word as irrelevant. In both cases, the following relevance weighing formula was used (Robertson & Jones, 1976):

$$\text{Relevance score} = \frac{\frac{r}{R-r}}{\frac{n-r}{N-R-n+r}},$$

where r stands for the hashtag's frequency of occurrence in relevant tweets, R stands for number of relevant tweets, n stands for total frequency of the hashtag, and N stands for total number of tweets. This formula examines the log-transformed ratio of the hashtag's occurrence in relevant tweets to its occurrence in irrelevant tweets. Higher scores thus indicate that the hashtag is more specific to relevant tweets. For both r and n, multiple occurrences of the hashtag within the same tweet were counted as 1. After manual filtering[1] of all the relevant hashtags that the two algorithms provided, we ended up with 177 relevant hashtags. To provide some examples, these hashtags included #Brusselsattack, #Brusselsbombing, #belgiumunderattack, #zaventem, #prayforbrussels, #prayforbelgium, and #istandwithbrussels.

Once we identified all the relevant hashtags, we searched all the tweets we collected via the Streaming API during the week following the attacks to identify users who tweeted or retweeted a tweet with a relevant hashtag, and extracted all the information about their influencers (i.e., the users they follow). We discovered 76,211 such users, who had a total of 101,455,590 influencers.[2] We then proceeded to extract the post-attack tweet history of these users and influencers with the REST API. As it was not feasible to extract this information for that many users and influencers, we decided to focus only on

users who define themselves as Dutch speakers and sampled 40% of them. We ensured that our sample had the same network size distribution as the entire Dutch sample. To do so, we divided the range of 0 to 3,000 influencers into steps of 50 and counted how many users in the Dutch sample have the number of influencers in each influencer's range. Then we sampled 40% of the users in each range. Our final sample had 959 unique users and 229,926 unique influencers. Using the REST API, we extracted the entire tweet history of all these users and influencers during the 2 weeks following the attack.

2.2. Coding

We tested the malleability of users' hashtag use by examining their overall hashtag use during the 2-weeks time period. In all analyses, we examined only original tweets by the user, not hashtags occurring in tweets they retweeted, since in retweets the user does not choose the hashtag that is included within the retweet. To test how malleable a user's hashtag use is overall, we counted how many different relevant hashtags a user used during the 2 weeks following the attack. In addition, we coded the number of unique relevant hashtags that each user has been exposed to during the 2-weeks time period (Hashtag Types), and the total number of relevant hashtags that the user has been exposed to (Hashtag Tokens). The difference between these two measures is that a hashtag that repeated 10 times in the user's feed would be counted once in the former measure, but 10 times in the latter measure.

2.3. Results

To test whether people who follow fewer Twitter users have more variable hashtag use, we ran a linear regression analysis on the number of unique hashtags the users had used during the 2 weeks period, with Number of Influencers, Hashtag Types, and Hashtags Tokens as predictors. Hashtag Types controls for the number of unique hashtags users have seen. If a user has only been exposed to five unique hashtags, he or she cannot adopt more than five, whereas a speaker who has been exposed to 15 unique hashtags can adopt many more. In theory, users can also invent new hashtags. Such cases were rare though. Indeed, the total number of hashtags a user adopted from his or her influencers and the total number of hashtags a user has used, regardless of whether any of the user's influencers had used them, correlated very highly ($r = .95$). Similarly, an analysis on the total number of hashtags that users used, regardless of whether they were used by their influencers, produces the same results as an analysis on the number of hashtags that users adopted from their influencers. We report here the former. Our analysis also included Hashtag Tokens to control for the fact that people might be more likely to use hashtags the more times they see them. The results show that, as predicted, people who follow more users use fewer different hashtags ($\beta = -1.05\text{e}-3$, $SE = 2.63\text{e}-4$, $t = -3.99$, $p < .001$; see Fig. 1). Additionally, the results show that people used more unique hashtags the more unique hashtags they have been exposed to ($\beta = 2.165\text{e}-2$,
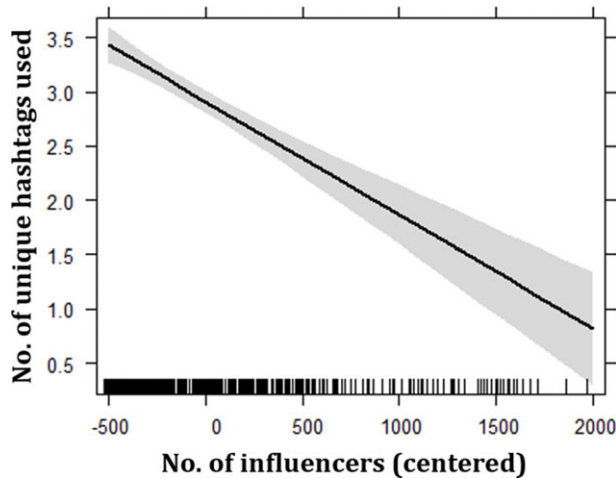
Fig. 1. The number of unique hashtags a user has used as dependent on the number of influencers he or she has.

$SE = 7.87e{-}3$, $t = 2.75$, $p < .01$), and the more times they have seen them ($\beta = 1.18e{-}3$, $SE = 2.73e{-}4$, $t = 4.33$, $p < .001$).[3]

The results of the study thus support the hypothesis and show that people who follow fewer Twitter users are more variable in their hashtag use. In particular, two users who are exposed to the same number of unique hashtags would differ in how many of them they would adopt depending on the number of people they follow. The more people they follow, the more they would be consistent in their use and the less they would adopt new hashtags they encounter.

3. General discussion

The goal of this study was to examine if the number of linguistic sources people have influences the malleability of their language use as reflected in the variability they exhibit. The results indicated that users who follow fewer others are more variable in their language use, and more likely to adopt new hashtags they encounter. They thus give further support to the hypothesis that the more speakers a person encounters, the less weight the person gives to the input from the speaker, and the less the person is influenced by the speaker. This finding is in line with previous experimental results that show that people with smaller social networks are more likely to change their phonological and lexical boundaries following exposure to patterns of speech that are different from their own. This study adds ecological validity to their previous results by showing that the same pattern extends to the real world when examining people's response to natural unmanipulated input.

One caveat is that we assumed that users saw the tweets that their influencers tweeted. Users, however, might not see all the tweets of their influencers but only a subset of them. One may even wonder whether users who follow more users might have more

tweets in their input and therefore might see fewer of them. At the same time, usually users who follow more users also spend longer on the website and read more tweets, making them even more likely to see all the tweets in their feed than those who follow fewer others. Unfortunately, at this point it is impossible to know which of these two patterns is stronger and which tweets the users indeed read. These results should therefore be taken with some caution and could be supported by further experimental studies which, similar to the previously reported lab studies, control and ensure that participants do not differ in their exposure to the target pattern.

Another potential caveat is that this study focused on tweets relating to a single event. Romero, Meeder, and Kleinberg (2011) have shown that the relationship between hashtag exposure and hashtag adoption depends on the topic. For example, in the political domain, repeated exposure to the same hashtag increases its adoption. This fits with the positive effect we found for Hashtag Tokens. In contrast, repeated exposure to hashtags in the music domain does not increase odds of adoption much. Future research should therefore examine whether the topic modulates the effect of network size on hashtag adoption, as the type of exposure that is most beneficial to spread as well as the weighing mechanism might depend on the topic.

These results could have implications for language change. One of the open questions in that field is how rare variants spread through the community. The results of this study suggest that people with smaller social networks might help in their propagation. Because people with smaller networks are more likely to adopt each variant they encounter, they could play an important role in the propagation of rare variants. One may wonder whether the role that people with smaller social networks could play is not limited precisely because of their small social network. That is, since people with smaller social networks interact with fewer others, they could only directly influence few others. Computational simulations, however, show that if we assume an inverse relationship between social network size and the weight ascribed to each source, as the results here and in previous studies show, then people with smaller social networks, despite their smaller direct circle of influence, are leaders of linguistic change (Lev-Ari, 2018). One may further wonder whether despite their theoretical ability to influence others, people with smaller social networks do not hold enough social capital to lead those in contact with them to learn from them and adopt their patterns, because they are likely to be non-central members of the community. Speakers' popularity could indeed modulate the degree to which people adopt their behavior. That said, much research has shown that people also adopt the behavior of non-central individuals. For example, it has been shown that weak ties are important for the propagation of trends across the community (e.g., Granovetter, 1973). Weak ties are not necessarily people with smaller social networks, but similarly to them, they are not central members in one's network. These studies thus show that non-central members of the community could be instrumental for the spread of trends, suggesting that the non-centrality of people with smaller social networks might not impede them from being instrumental for the spread of linguistic phenomena. That said, further work is required to assess whether people with smaller social networks are indeed propagators of linguistic change.

To conclude, this study provides further support to the hypothesis that people with smaller social networks are more malleable and therefore variable in their language use. This study examines natural and unmanipulated language use, thus indicating that the previous results that were obtained under well-controlled laboratory situations extend to the spread of linguistic patterns in the real world. These results thus show how properties of our social environment influence how we learn, use, and potentially propagate linguistic patterns.

Notes

1. Manual filtering was required because, by their nature, the relevance algorithms are only approximate. Furthermore, we opted to use particularly broad relevance algorithm to maximize the odds of finding all relevant hashtags, including rare ones. This necessarily led to the ranking of irrelevant hashtags as well. For example, the time-based relevance algorithm examines the increase in hashtag appearance on the day of the event, but it is not restricted by topic. Thus, hashtags referring to other events that occurred around that day were ranked high as well.
2. This number is an overestimate because some people might follow the same users, but the overlap is not as common as one would assume, and the final number of unique users is still in the tens of millions.
3. One may wonder whether this effect is due to differences in the entropy in the input that users with different number of influencers have. Examination of the data showed that Number of Influencers *positively* correlates with entropy ($r = .45$, $p < .0001$). We reran the model with the inclusion of entropy. Entropy did not have an effect ($t = -1.3$, $p > .1$), and all other effects remained statistically significant and in the same direction. In particular, Number of Influencers still had a negative effect on number of used hashtags ($\beta = -0.1e-02$, $SE = 0.26e-3$, $t = -4.025$, $p = 6.58e-05$).

References

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological Science*, *20*, 372–377.

Dale, R., & Lupyan, G. (2012). Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems*, *15* (03n04), 1150017.

Forster, K. I. (1976). Accessing the internal lexicon. In R. J. Wales & E. C. T. Walker (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627–635.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105* (2), 251–279.

Grainger, J., O'regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception and Psychophysics*, *45*(3), 189–195.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Haun, D. B., Rekers, Y., & Tomasello, M. (2012). Majority-biased transmission in chimpanzees and human children, but not orangutans. *Current Biology*, *22*(8), 727–731.

Herrmann, P., Legare, C. H., Harris, P. L., & Whitehouse, H. (2013). Stick to the script: The effect of witnessing multiple actors on children's imitation. *Cognition*, *129*, 536–543.

Hills, T. T., & Adelman, J. S. (2015). Recent evolution of learnability in American English from 1800 to 2000. *Cognition*, *143*, 87–92.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego, CA: Academic Press.

Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. In E. W. Hinrichs & D. Roth (Eds.), *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics* (vol. 1, pp. 423–430). Sapporo, Japan.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203.

Lev-Ari, S. (2017). Talking to fewer people leads to having more malleable linguistic representations. *PLoS ONE*, *12*(8), e0183593.

Lev-Ari, S. (2018). Social network size can influence linguistic malleability and the propagation of linguistic change. Cognition *176*, 31–39. https://doi.org/10.1016/j.cognition.2018.03.003.

Lev-Ari, S., & Shao, Z. (2017). How social network heterogeneity facilitates lexical access and lexical prediction. *Memory & Cognition*, *45*(3), 528–538.

Luce, P. A. (1986). Neighborhoods of words in the mental lexicon. Research on Speech Perception, Technical Report No. 6.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014). The Stanford coreNLP natural language processing toolkit. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, System Demonstrations* (pp. 55–60). Association for Computational Linguistics: Stroudsburg, PA.

Nettle, D. (1999). Using social impact theory to simulate language change. *Lingua*, *108*(2), 95–117.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238.

Robertson, S. E., & Jones, K. S. (1976). Relevance weighting of search terms. *Journal of the Association for Information Science and Technology*, *27*(3), 129–146.

Romero, D. M., Meeder, B., & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th International Conference on World Wide Web* (pp. 695–704). Hyderabad, India: ACM.

Roodenrys, S., Hulme, C., Lethbridge, A., Hinton, M., & Nimmo, L. M. (2002). Word-frequency and phonological-neighborhood effects on verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1019–1034.

Schillaci, R. S., & Kelemen, D. (2013). Children's conformity when acquiring novel conventions: The case of artifacts. *Journal of Cognition and Development*, *15*, 569–583.

Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(1), 28–39.