Title

Constraints on Learning Non-Adjacent Dependencies (NADs) of Visual Stimuli

Permalink

https://escholarship.org/uc/item/1b75k23m

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors

Li, Jia
Mintz, Toben

Publication Date

2015

Constraints on Learning Non-Adjacent Dependencies (NADs) of Visual Stimuli

Jia Li (JIA10@Usc.Edu)
Department of Psychology, SGM 501, 3620 S.McClintock Avenue
Los Angeles, CA 90089-1061 USA

Toben Mintz (TMINTZ@Usc.Edu)
Department of Psychology, SGM 501, 3620 S.McClintock Avenue
Los Angeles, CA 90089-1061 USA
Department of Psychology, GFS 301, 3601 Watt Way
Los Angeles, CA 90089-1693 USA

Abstract

Non-adjacent dependencies (NADs) refer to dependencies between items that are not adjacent in a sequence. Peña et al. (2002) discovered adult participants could learn the NADs of syllables in an artificial language when there were 25ms pauses before and after the NADs. Studies using videos of human body movements showed similar learning outcomes (Endress & Wood, 2011). However, participants failed to learn the NADs with respect to non-linguistic acoustic stimuli, such as tones or noises (Gebhart, Newport, & Aslin, 2009). Four experiments in this study examined the constraints on learning the NADs of visual stimuli. We propose that acquisition of the NADs requires the sequences be packed into a coherent unit, and the motor system provides the require packaging for stimuli that can be mapped onto motor representation. Implications on the acquisition of syllable NADs are discussed.

Keywords: statistical learning; Non-Adjacent Dependencies (NADs); sequence learning; visual stimuli

Introduction

Our sensory experience is full of regularities distributed over time. How do we track and discover these regularities quickly and unintentionally? Research on statistical learning has shown that humans can discover both visual and auditory regularities by tracking the co-occurrence patterns of the stimuli. Studies on statistical learning of temporal regularities can be broadly divided into two categories based on the types of distributional cues: (1) Adjacent dependencies, where cues occur among temporally adjacent stimuli, and (2) Non-adjacent dependencies (NADs), where cues are interspersed over time. An example of the first category is human infants' ability to track distributional cues—transitional probabilities (TPs)—between syllables in a speech stream (Saffran, Aslin, & Newport, 1996). An example of the second category is our ability to track dependencies such as agreement patterns (e.g., *is* sleep*ing*; Santlemann & Juszcyk, 1998). Previous research has confirmed human subjects' consistently similar capacity to learn adjacent dependencies in both linguistic and non-linguistic stimuli, such as tones, noises, images, and body movements (Creel, Newport, & Aslin, 2004; Endress & Wood, 2011; Fiser & Aslin, 2002; Gebhart, Newport, & Aslin, 2009; Kirkham, Slemmer, & Johnson, 2002; Saffran et al., 1999; Turk-Browne, & Scholl, 2009). By contrast, the acquisition of NADs exhibits differing characteristics with different types of stimuli.

Statistical Learning of NADs

Studies have shown that participants can acquire NADs of tones, phonemes and syllables, with certain limitations.

In studies of the acquisition of NADs among phonemes, phonemes with NADs were either all consonants (e.g., t_k_p_, with "_" indicating spaces for vowels), or all vowels (e.g. _e_i_u, with "_" indicating spaces for consonants) (Bonatti, Peña, Nespor, & Mehler, 2005; Newport & Aslin, 2004). Newport and Aslin (2004) proposed that consonants and vowels are segmented into different phonological tiers, as proposed by Autosegmental phonology (Goldsmith, 1976). Therefore learning the NADs between consonants or vowels equates to learning adjacent dependencies between them within their respective tier.

Peña et al. (2002) tested whether participants were able to acquire the NADs of syllables such as pu__ki in an artificial language …pulikibedugapuraki…. They found that adult participants were able to learn NADs between syllables when there were 25ms pauses before and after the NADs, but failed to do so absent the brief pauses.

With respect to non-linguistic stimuli, studies with tones (Creel et al., 2004; Gebhart et al., 2009), noises (Gebhart et al., 2009), and abstract images (Turk-Browne, Jungé, & Scholl, 2005) indicated that the NADs of these non-linguistic stimuli can be readily learned only when the units with NADs are perceptually similar, following Gestalt's principles of perception (Wertheimer, 1923). For example, Creel et al. (2004) showed that participants successfully discriminated tone triplets that were interleaved with other tones in a sequence of tones, when the tone triplets had distinctive pitches (from separate octaves). When the perceptual cues were removed from the stimuli, studies using tones, noises (Gebhart et al., 2009), and abstract images (Li & Mintz, 2014) have failed to find evidence that the NADs of these acoustic or visual stimuli could be readily learned, even when the NADs were bracketed with pauses as in Peña et al. (2002). However, Endress and Wood (2011) tested the acquisition of NADs using videos of human body movements, and observed learning effects

similar to those observed in studies using syllables. Further, such learning did not depend on the viewing angle of the stimuli (Endress & Wood, 2011).

Taken together, the above studies suggest that: (1) Perceptual cues facilitate the acquisition of NADs; (2) pauses between sequences with NADs also facilitate the acquisition of NADs, but only with respect to syllables and human body movements, not with respect to non-linguistic acoustic stimuli, such as tones or noises.

Mechanisms Underlying Acquisition of NADs

The above analysis gave rise to two questions. First, what role did the pauses play in the acquisition of the NADs of stimuli? Second, why did pauses facilitate the acquisition of the NADs of certain types of stimuli (syllables and body movements) but not of others (abstract images and tones)?

One possible answer to the first question is that the pauses bracketed syllable sequences; as a result, syllables at the beginning and ending positions occupied special edge positions. Henson's Start-End Model (SEM) proposed that, the representation system places a "start marker" and an "end marker" in each sequence (Henson, 1998, 1999) and that the items' positions are recorded as their distance to one of the two markers. Building on this idea, the dual mechanisms account (Endress & Bonatti, 2007; Endress, Nespor, & Mehler, 2009; Endress & Mehler, 2009) proposed a positional learning mechanism that rapidly recording syllables' positions relative to the edges of the sequences during statistical learning of syllable sequences in Peña et al. (2002). The positional learning mechanism plausibly explains the statistical learning of syllable NADs.

So far, no compelling answer has been proposed to the second question regarding the reason why pauses facilitate the acquisition of NADs of syllables and body movements, and not tons or images. This naturally invites speculation that learning syllables and body movements is governed by the same underlying mechanism. If it is, that leads to a puzzle: what kind of learning mechanism would be engaged by speech and body movements, but not by other stimuli.

One possibility is that syllables and body movements fluidly transform from one stimulus to the next (unlike distinct tones and images which shift sharply from one stimulus to the next), which facilitates the acquisition of the NADs. Sensitivity to the NADs of movements is not particular to human body movements, and the NADs of any movements at the beginning and ending of a continuous sequence of motion can be learned. It has been shown that human perception is generally sensitive to dynamics of an agent. In an object recognition task in Vuong & Tarr (2004), participants were first presented with a rotating object, and then with a single view of the object, and were then asked to indicate whether the test object was the same rotating object from the training. Participants responded faster and more accurately when the test views were from the beginning or the end of the rotation. Participants were even sensitive to unattested views that preceded or followed the trajectory of rotation in the training. In Vuong & Tarr (2004), each object

rotated in a single direction, but it is equally possible that higher familiarity with the particular movements at the beginning and ending of a continuous series of movements would also results in the acquisition of NADs.

Another possibility is that the acquisition of NADs of speech and that of body movements share common cognitive processes, given that syllable sequences may be perceived as sequences of corresponding vocal movements. For example, the motor theory of speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985) posits that in perceiving speech, human beings map the acoustic signal to articulatory gestures. Likewise, the representation of visually perceived human movements involves the activation of motor representations by the perceiver (e.g., Wilson & Knoblich, 2005). Thus, it is possible that a kind of motor sequence learning is the common underlying mechanism supporting the statistical learning of syllables and body movements. The beginning and ending of the motor sequences are prominent since they mark the change of status, from stillness to motion and from motion to stillness. This might facilitate the detection of dependency patterns in which the beginnings and endings take part.

The current study examines each of these two possibilities as potential explanations for the mechanism underlying NAD acquisition. It is worth noting that the two explanations are not mutually exclusive. It is possible that both play a function in bracketing sequences, and thus jointly contribute to the acquisition of NADs.

Study Synopsis

In the current study, we first ask whether the NADs of non-human movements can be acquired, by replicating the Endress & Wood (2011) findings regarding human movements. Experiment 2 tests subjects' acquisition of NADs from a of objects moving in a manner that would be biologically impossible for human beings, but is nonetheless continuous and coherent. Next, the study tests if continuous movement is critical for the acquisition of NADs. Experiment 3 tests NAD learning with sequences of static images of body postures, and Experiment 4 tests NAD learning with static images of objects. Experiments 1 and 3 use stimuli that can be mapped onto representation of body movements while Experiments 2 and 4 do not. Thus, the four experiments investigate the continuous movement hypothesis and the motor sequence learning hypothesis.

Experiment 1: NADs of Body Movements (Replication of Endress and Wood (2011)

Experiment 1 replicated the finding in Endress and Wood (2011) that adult participants were capable of acquiring NADs of human body movements.

Methods

Participants Twenty undergraduate students from the University of Southern California (USC) were recruited

from the USC Psychology Subject Pool. Their participation in the experiment was compensated with course credits.

Apparatus and Stimuli The original episodes of body movements from Endress and Wood (2011) were used to create training and testing stimuli in this experiment. In each original episode, an animated male agent performed a movement (e.g., bending), and the movements started and ended in the same neutral, still, standing position, referred to herein as the "neutral position" consistent with Endress and Wood (2011). There were two major differences between the stimuli in the current experiment and in Endress and Wood (2011). First, the pauses between sequences in this experiment were a blank screen, instead of the neutral position in Endress and Wood (2011), due to the fact that neutral positions could not be used as intervals in Experiments 3 and 4, and the study sought to minimize the differences between the designs of the experiments. Second, the parameters of the visual presentation were different. In this experiment, each movement episode lasted 625ms with 15 frames presented at a frame rate of 24 frames/second. Each frame was sized 480×468 pixels.

Figure 1. Frames excerpted from the body movement animations used in Experiment 1 (depicting the maximum extent of movement), which were also the still images used in Experiment 3. The stimuli are the original stimuli used in Endress and Wood (2011).

Training The structure of the training and testing stimuli was similar to that in Peña et al. (2002), with nine syllables replaced with nine body movements. Nine triplets were created by pairing each of three pairs of NADs (a_b, c_d, e_f, with each letter representing a body movement) with each of three middle items (x,y,z), and 20 repetitions of the nine triplets were randomly concatenated into a continuous visual stream. We imposed a sequencing constraint such that each triplet was immediately followed by a triplet with a different NAD, and a different middle item. Each triplet was presented for 1875ms, with a 125ms pause between the triplets, resulting in an entire training sequences of 6'22''.

Testing After exposure to the training set, participants were tested on their preference for two kinds of triplets: (1) Rule-Triplets: Three-item sequences with the correct NADs paired with middle items that were unattested during training in the particular NAD (e.g., acb, cfd), and (2) Part-Triplets: Three-item sequences spanning two consecutive,

attested triplets (e.g., xbc, dez). Rule-triplets and part-triplets differed in two major ways: 1) participants were actually exposed to part-triplets during training, but not to rule-triplets; 2) rule-triplets contained the same NADs as trained triplets, while part-triplets did not. There were 36 test pairs contrasting rule- and part-triplets. The presentation order of the two types of sequences within a pair and the response buttons were counterbalanced.

Results and Discussion
Participants' responses were coded as binary variables, with preference for rule-triplets coded as 1 and preference for part-triplets coded as 0. A logistic regression model was used to compare participants' choice with the chance level (0.5), with the binary responses as the dependent variable; the model controlled for variance based on participants and test questions. Participants' mean preference for rule-triplets over 36 testing pairs, the standard deviation, the intercept (β), z, and p-value from the logistic regression, are listed in Table 1. Intercept (β) indicates deviation from the chance level in the choice tests between rule-triplets and part-triplets. In Experiment 1, out of 36 test trials, the average number of trials in which participants preferred rule-triplets is 22.75 (SD = 6.09), approximately 63.19%. Logistic regression of Experiment 1 yielded a significant intercept (ß = 0.61, SE = 0.18, z = 3.37, p < .001), indicating participants considered rule-triplets to be more familiar. Figure 2 shows each participant's percentage of preference for rule-triplets in this experiment. The above analysis was done using R 3.0.2 GUI 1.62 Snow Leopard build (6558), and lme4 R package, version 1.0-4, and graphed with graphics version 3.0.2.

This experiment confirmed the findings in Endress and Wood (2011) that participants were sensitive to the NADs of body movements. Experiment 2 will examine if participants would be similarly capable of learning NADs of movements performed by a non-human object.

Table 1: Mean preference for rule-triplets over 36 trials including standard deviation, **the intercept (β), z, and p-value from the logistic regression of** each experiment.

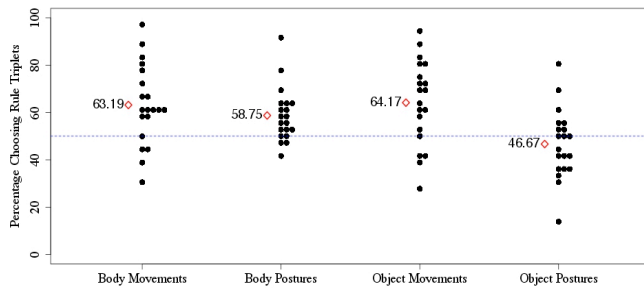| Exp	Mean preference to rule-triplets (SD)	Intercept (β) (SE)	z	p
1.	22.75 (6.09)	0.61 (0.18)	3.37	<.001
2.	23.1 (6.45)	0.66 (0.19)	3.49	<.001
3.	21.15 (4.12)	0.64 (0.11)	3.41	<.001
4.	16.8 (5.27)	-0.14 (0.14)	-1.04	0.3

Figure 2: Results of four experiments. Each dot represents percentage of preference for rule-triplets for each participant, and the diamond represents group mean. The Dotted line indicates chance level (50%).

Experiment 2: NADs of Object Movements and Transformations

Experiment 2 explored if the NADs of object movements and transformations that could not be plausibly performed by human agents could also be acquired.

Methods

The methods were the same as in Experiment 1, except that each body movement was replaced with an animated object movement or transformation, as shown in Figure 3. The red blanket-shaped object (as in the cell titled "Neutral Position") performed movements that cannot be mapped onto human motor representations. The videos of object transformations were created using 3ds Max.
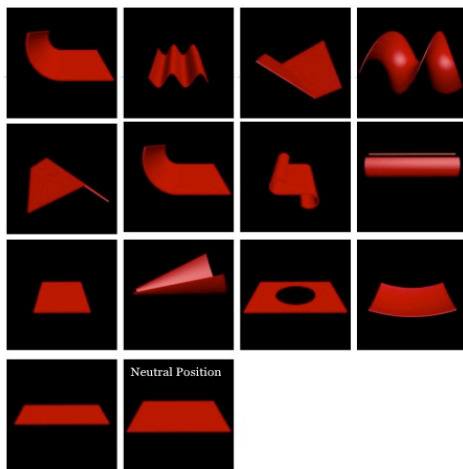

Figure 3. Depiction of the object transformations and neutral position used in Experiment 2, and the images used in Experiment 4. In Experiment 2, movement into each position was continuous from the flat, neutral position.

Results and Discussion

In Experiment 2, the average preference for rule-triplets was 23.1 (SD = 6.45) over 36 test trials, which is about 64.17% (See Table 1 and Figure 2). Analysis of Experiment 2 yielded a significant intercept (ß = 0.66, SE = 0.19, z = 3.49,

p < .001), suggesting participants considered rule-triplets more familiar.

This result suggests that continuous movements aid the acquisition of NADs regardless of whether the movements can be mapped to motor representations. Bracketed continuous movements provide sufficient packaging of sequences to facilitate learning patterns involving the sequence beginning and end. The following two studies further examine this idea by testing whether participants are equally likely to learn the NADs of static images of body postures (Experiment 3) and object postures (Experiment 4), rather than continuous movements. It is possible that they would fail to do so in both Experiment 3 and 4 due to the lack of continuous movement. It is also possible they would succeed or fail in both experiments. Another possibility is that participants would fail in Experiment 4, but succeed in Experiment 3, because the images of body postures provide sequences of implied body actions that are perceived as continuous movement (Shiffrar & Freyd, 1990; Urgesi, Moro, Candidi, & Aglioti, 2006), which in turn triggers motor learning of continuous movements from one posture to the next, as in Experiment 1. To minimize the differences between the stimuli of Experiments 3 and 4 and the stimuli of Experiments 1 and 2, the 9th frames of the movement episodes in Experiments 1 and 2 were used as the image stimuli.

Experiment 3: NADs of Body Postures

Experiment 3 examined if participants can acquire the NADs of the images of human body postures as they did in Experiment 1.

Methods

The methods were the same as in Experiment 1, except that each body movement was replaced with the 9th frame of the video of each movement. Each image sized 480×468 pixels. Each image was presented for 625ms, same as the duration of each movement video in Experiments 1 and 2. The between-triplet pauses were also 125ms. The total duration of the training sequence was 6'22''.

Results and Discussion

The average preference for rule-triplets in Experiment 3 was 21.15 (SD = 4.12), 58.75% (See Table 1 and Figure 2). Logistic regression yielded a significant intercept ß = 0.64, SE = 0.11, z = 3.41, p < .001, indicating that participants were sensitive to the NADs of body postures.

Still, it is unknown if these findings would extend to images in general, or if participants are sensitive to body postures because viewers interpret the posture sequences as continuous movement, which then functions as in Experiment 1. Experiment 4 tested whether participants could also acquire the NADs of images of the objects in different postures.

Experiment 4: NADs of Object Postures

Experiment 4 explored if participants can acquire the NADs of object postures as they did with human postures in Experiment 3.

Methods

The methods were the same as in Experiment 3, except that each body posture was replaced with a static image of an object posture, which was the 9th frame of the corresponding video in Experiment 2.

Results and Discussion

In Experiment 4, the average preference for rule-triplets was 16.8 (SD = 5.27), around 46.67% (see Table 1 and Figure 2). The intercept from logistic regression was not statistically significant, (ß = -0.14, SE = 0.14, z = 1.04, p = 0.3), suggesting that participants failed to distinguish between rule-triplets and part-triplets,.

The outcome that participants successfully learned the NADs of static body postures, but not of static object postures, suggests two points. First, with simple objects, as opposed to human postures, packaging the sequences through continuous movement appears to be necessary for acquiring NADs. Absent such packaging, participants failed to learn the NADs of the images. Second, processing and representation of body movement sequences appears to be special.

Discussion

The inquiry into the representation of body movements and continuous object movements stemmed from comparable results obtained in studies of the NADs of syllables (Peña et al., 2002) and body movements (Endress & Wood, 2001). Both studies suggested that dependency rules involving NADs of syllables and body movements could be learned, with the condition that the sequences with NADs were bracketed by pauses (Peña et al., 2002). The four experiments in the current study probed two questions regarding visual statistical learning of NADs: (1) if such learning pertains only to stimuli in the form of human movements; (2) if continuous movement has an impact on participants' acquisition of the NADs of non-human object stimuli. Experiment 1 replicated one major finding in Endress and Wood (2011) and confirmed that participants could acquire NADs of body movements under the current experimental conditions. Experiment 2 probed if movements' susceptibility to being mapped onto the human body was a necessary condition for learning their NADs, by replacing the human agent with an object performing movements that could not be represented as human body movements, and found that participants similarly learned the NADs of the object movements. Experiments 3 and 4 used static images that depicted the maximal extent of the movements depicted in Experiments 1 and 2.

Experiment 1 (with body movements) and 2 (with object movements) demonstrated the general capacity of the human cognitive system to track and learn the beginning and ending movements of a continuous sequence of movements, regardless of the agent performing them, or whether the movements were human-like. However, this does not mean that the representation and processing of movements were the same for body movements and object movements/transformations. The differing results of Experiment 2 (with static human postures) and Experiment 4 (with static object postures) indicated differences in the underlying processing of static images.

In Experiment 3, participants successfully learned the NADs of the static body postures of the same agent in Experiment 1. While the results could be explained by a separate representation system for object sequences and body postures, the contrasting results suggest the involvement of the motor system in visual sequence learning. The discrete images of static body postures, once mapped onto a representation of the observer's motor system, are perceived as continuous body movements (Shiffrar & Freyd, 1990; Urgesi, et al., 2006). This may in turn activate motor representations similar to those that are activated when continuous motions were viewed, thereby achieving the same packaging of the movement-triples that highlights the beginnings and ends, leading to successful learning of NADs, as in Experiment 1. In other words, the motor system facilitates the linkage of distinct body postures into coherent movements. In fact, viewers of the static body posture sequences themselves reported that the sequences created a sense of continuous movement. With respect to object postures, since they cannot be mapped onto the motor system, they are still perceived as separate images of an object. Therefore, participants failed to learn the NADs of these different objects.

The facilitating role of the motor system in statistical learning of visual stimuli has implications for understanding the underlying mechanisms of statistical learning in the domain of language. With acoustic stimuli, acquisition of NADs has been observed with syllables (Peña et al., 2002) but not tones or noises (Gebhart et al., 2009; Li and Mintz, 2014). The Motor Theory of speech perception proposes that the perception of syllables is mapped onto vocal gestures, and those gesture representations drive perception (Liberman et al., 1967; Liberman & Mattingly, 1985). More contemporary research also implicates motor representations in the perception of other individuals' movements, in speech, and more broadly (Fadig, Craighero, & Olivier, 2005; Skipper, Nusbaum, & Small, 2005). If this is so, the syllable sequences could be represented as sequences of vocal movements by the motor system. Learning syllable sequences would then boil down to motor sequence learning, similar to learning body movements and static body postures. Therefore, we propose that motor sequence learning is a critical part of learning syllable dependency patterns in speech, in that it provides a kind of packaging of sequences that highlights beginnings and ends (Henson, 1998, 1999), and therefore facilitates the learning of the

patterns between them. We are currently planning studies to directly test this hypothesis.

Taken together, these experiments support the hypothesis that learning non-adjacent dependencies (NADs) requires bracketing of sequences. Moreover, we propose that learning NADs requires that the sequence in question be packaged into a coherent unit. We further show that the motor system can provide the required packaging, and NADs can be learned when the stimuli can be mapped onto motor representations.

References

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: the role of consonants and vowels in continuous speech processing. *Psychological Science, 16*(6), 451–459. doi:10.1111/j.0956-7976.2005.01556.x

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1119–1130. doi:10.1037/0278-7393.30.5.1119

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition, 105*(2), 247–299.

Endress, A. D., Nespor, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, *13*(8), 348–353. doi:10.1016/j.tics.2009.05.005

Endress, A. D., & Mehler, J. (2009). Primitive computations in speech processing. *Quarterly Journal of Experimental Psychology, 62*(11), 2187–209.

Endress, A. D., & Wood, J. N. (2011). From movements to actions: two mechanisms for learning action sequences. *Cognitive Psychology*, *63*(3), 141–71.

Fadiga, L., Craighero, L., & Olivier, E. (2005). Human motor cortex excitability during the perception of others' action. *Current opinion in neurobiology*, *15*(2), 213-218.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 458–467. doi:10.1037//0278-7393.28.3.458

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review, 16*(3), 486–490.

Henson, R. N. A. (1998). Short-term memory for serial order: the Start-End Model. *Cognitive Psychology, 36*(2), 73–137. doi:10.1006/cogp.1998.0685

Henson, R. N. A. (1999). Positional information in short-term memory: Relative or absolute? *Memory & Cognition*, *27*(5), 915–927.

Kirkham, N. Z., Slemmer, J. A, & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–42.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*(5593), 604–7. doi:10.1126/science.1072901.

Li, J., & Mintz, T. H. (2014). Learning Non-Adjacent Dependencies of Tones and Images. Arizona Linguistic Circle 8, Tucson, USA.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74(6)*, 431-461.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127–162. doi:10.1016/S0010-0285(03)00128-2

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52.

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition*, *69*(2), 105–134.

Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science*, *1*(4), 257–264. doi:10.1111/j.1467-9280.1990.tb00210.x

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage*, *25*(1), 76-89.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–64. doi:10.1037/0096-3445.134.4.552

Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 195–202.

Urgesi, C., Moro, V., Candidi, M., & Aglioti, S. M. (2006). Mapping implied body actions in the human motor system. *Journal of Neuroscience*, *26*(30), 7942–7949. doi:10.1523/JNEUROSCI.1289-06.2006

Vuong, Q, C. & Tarr, M.J. (2004). Rotation direction affects object recognition. *Visual Resaerch, 44*, 1717-1730.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460–473. doi:10.1037/0033-2909.131.3.460