# BREATHING POSITION AND THE PERCEPTION OF CONSONANTS

**Henny Yeung [*1,3], Mark Scott [†2] and Lionel Granjon [‡3]**
[1]Department of Linguistics, Simon Fraser University, Burnaby, BC, Canada
[2]Department of English Literature & Linguistics, Qatar University, Doha, Qatar
[3]Laboratoire Psychologie de la Perception, CNRS & Université Paris Descartes, Paris, France

## 1 Introduction

Several theories of speech perception and acquisition suggest a role for the motor system in auditory processing. One class of theories suggests that motor information plays an especially important role when the perceived speech is noisy, ambiguous, or otherwise hard to process, and that the motor system can be used to make sensory predictions about the incoming auditory signal to fill in missing information [1] (see also [2]). A source of evidence for these theories comes from studies that require experimental participants to move their vocal-tract articulators in speech-like ways while perceiving ambiguous or noisy speech [3]–[6]: Results show that these movements influence the perception of speech signal in ways that other non-auditory sources of complementary speech information can influence speech perception, like the McGurk effect (i.e., seeing visual speech changes the perception of auditory speech [7]).

It remains unclear what precise aspects of speech planning and production generate sensory predictions, which might influence auditory speech perception. Broadly speaking, previous reports suggest two distinct sources of information. First, speech perception seems to be modulated by the internal rehearsal or planning of speech gestures (inner speech), even if no explicit movements are made [3], [8]. Second, peripheral sensory information about the actual dynamic movements of speech articulators seems to result in a similar effect. For example, when there is an external source eliciting sensory feedback from skin receptors—a robot deforming facial skin in speech-like ways—speech perception is modulated, even if perceivers are not explicitly engaging in any speech rehearsal or planning [9].

Several questions about these sensorimotor influences on perception persist. In the present report, we focus on the specificity of articulatory information in these pathways, asking what kinds of motor activity are sufficient to trigger sensory predictions, and thus influence speech perception. We considered an extreme case in this study, as participants were asked just to maintain a static vocal tract position (i.e., breathing either through the nose or through the mouth) while perceiving speech. This tested the idea that speech-related motor-induced sensory predictions are elicited in an automatic, pervasive, and broad manner, perhaps even when simply breathing in a particular way (i.e., having vocal tract configurations that only loosely match real speech gestures).

* henny_yeung@sfu.ca
† mark.a.j.scott@gmail.com
‡ lionel.granjon@parisdescartes.fr

## 2 Methods

Experimental participants performed a speech identification task in two conditions. The Experimental Condition involved a classification of an /ada/-/ana/ continuum, while the Control Condition involved classification of a /ada/-/aga/ continuum. In coarse terms, continua endpoints differed either in the position of the velum (up or down for /ada/ and /ana/, respectively), or the position of the tongue tip (raised or lowered for /ada/ and /aga/, respectively).

The critical manipulation involved instructions to the participants to either breathe through the mouth or nose while performing the task. This necessarily changes the position of the velum (up or down), and we predicted that breathing position would have an effect on identifications in the Experimental Condition, but not the Control Condition.

### 2.1 Participants

Forty-nine native monolingual French speakers participated. All were between the ages of 18-30 with normal hearing.

### 2.2 Stimuli

A female French native speaker was recorded saying /ada/, /ana/, and /aga/. One token of each disyllable was then used as the endpoints to create the two continua, each with 10,000 steps, using the STRAIGHT software package [10].

### 2.3 Procedure

Participants were tested using a laptop computer with headphones in a quiet room. In total, there were 8 experimental blocks presented in a counterbalanced order: 2 breathing positions (Nose and Mouth) x 2 conditions (Experimental and Control) x 2 block repetitions.

Within each block, a double staircase procedure was used to sample participants' perceptual boundaries [11]. A "high" staircase started at token 7600, which was closer to either the /ana/ or /aga/ continuum endpoints, while a "low" staircase started at token 2400, which was at the /ada/ endpoints of the continua. Sixteen staircase steps of decreasing sizes were used, and trials alternated randomly between either the high or low staircase.

In each trial one token was presented, after which a button was pressed to indicate which continuum endpoint was heard. A new trial began after a 150ms delay.

## 3 Results

### 3.1 Analysis

A mixed-effects logistic regression was used to model responses on each continua separately, with the expectation that there would be an effect of breathing place in the Experimental Condition (the /ada/-/ana/ continuum), but not in the Control Condition (the /ada/-/aga/ continuum).

For fixed effects, breathing place (Mouth or Nose) and continuum token (coded as 0 to 100) were both entered into the model. By-subjects random slopes were entered for breathing place, continuum token, and staircase (four different staircases were run for each condition: high and low staircases in each of the two block repetitions). All analyses were conducted using the *lme4* package in R [12].

### 3.2 Experimental Condition (/ada/-/ana/)

Results of the mixed effects logistic regression suggested that both fixed factors (breathing place and continuum token) influenced the likelihood of participants' category response. For breathing place, switching from mouth breathing to nose breathing significantly affected response probability ($\chi2(1) = 8.48$, $p = 0.0036$), and the model estimate (Table 1) suggested an increase in the probability of /ana/ responses when breathing through the nose.

As predicted, when presenting tokens from higher on the continuum (more /ana/-like), responses were also more likely to be perceived as /ana/ ($\chi2(1) = 83.45$, $p < 0.001$).

| | Estimate | S.E. |
|---|---|---|
| Intercept | -3.43 | .39 |
| Breathing Place | **.34** | **.11** |
| Continuum Token | **.12** | **.01** |

**Table 1:** Fixed effects in the Experimental Condition (bold font indicates a significant fixed factor at α = .01).

### 3.3 Control Condition (/ada/-/aga/)

Results of the mixed effects logistic regression suggested that only one of the fixed factors (continuum token) influenced the likelihood of participants' category response. For breathing place, switching from mouth breathing to nose breathing did not significantly affect the response probability ($\chi2(1) = 1.91$, $p = 0.17$), and as can be seen in Table 1, the standard error of the model's coefficient indicated no consistent pattern of breathing place.

As predicted, when presenting tokens from higher on the continuum (more /aga/-like), responses were also more likely to be perceived as /aga/ ($\chi2(1) = 24.10$, $p < 0.001$).

| | Estimate | S.E. |
|---|---|---|
| Intercept | -2.08 | .45 |
| Breathing Place | .17 | .12 |
| Continuum Token | **.044** | **.008** |

**Table 2:** Fixed effects in the Control Condition (bold font indicates a significant fixed factor at α = .01).

## 4 Discussion

Results suggested that simply instructing experimental participants to breathe through the nose biased speech identifications towards the /ana/ category on the /ada/-/ana/ continuum in the Experimental Condition. No such effect was found for the /ada/-/aga/ continuum in the Control Condition.

## 5 Conclusion

These results are striking, as they show the effect of highly subtle motor information on auditory speech identification. In other words, simply instructing participants to breathe through their mouth or nose (and thus changing the position of the velum) can affect perceptual judgments of ambiguous speech sounds (between /ada/ and /ana/). Results extend previous studies in showing that speech identification is influenced by motor planning not linked to auditory imagery, and by motor movements that are not speech-like.

Together, results support speech-processing theories that incorporate sensory predictions from the motor system [1], [2]. The present data suggest that such sensory predictions are highly automatic and pervasive, generated from even very basic information about static positions of oral articulators.

## References

[1] J. I. Skipper, V. van Wassenhove, H. C. Nusbaum, and S. L. Small. Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cereb. Cortex*., 17 :10, 2007.

[2] M. J. Pickering and S. Garrod. An integrated theory of language production and comprehension. *Behav. Brain Sci.*, 36 :4, 2013.

[3] M. Scott, H. H. Yeung, B. W. Gick, and J. F. Werker. Inner speech captures the perception of external speech. *J. Acoust. Soc. Am.*, 133 :4, 2013.

[4] M. Sato, E. Troille, L. Ménard, M.-A. Cathiard, and V. Gracco. Silent articulation modulates auditory and audiovisual speech perception. *Exp. Brain Res.*, 227, 2013.

[5] M. Sams, R. Möttönen, and T. Sihvonen. Seeing and hearing others and oneself talk. *Cogn. Brain Res.,* 23 :2–3, 2005.

[6] T. Mochida, T. Kimura, S. Hiroya, N. Kitagawa, H. Gomi, and T. Kondo. Speech misperception: speaking and seeing interfere differently with hearing. *PLoS One*, 8 :7, 2013.

[7] H. McGurk and J. MacDonald. Hearing lips and seeing voices. *Nature*, 264 :5588, 1976.

[8] M. Scott. Corollary discharge provides the sensory content of inner speech. *Psychol. Sci.*, 24 :9, 2013.

[9] T. Ito, M. Tiede, and D. J. Ostry. Somatosensory function in speech perception. *Proc. Natl. Acad. Sci. USA*, 106 :4, 2009.

[10] H. Kawahara, T. Irino, and M. Morise. An interference-free representation of instantaneous frequency of periodic signals and its application to F0 extraction. *ICASSP Proceedings*, 2011, pp. 5420–5423.

[11] T. N. Cornsweet. The staircase-method in psychophysics. *Am. J. Psychol*., 75, 1962.

[12] D. Bates, M. Mächler, B. M. Bolker, and S. C. Walker. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 :1, 2015.