Seeing voices: the role of multimodal cues in vocal learning
Varkevisser, J.M.

Citation
Varkevisser, J. M. (2022, October 20). *Seeing voices: the role of multimodal cues in vocal learning*. Retrieved from https://hdl.handle.net/1887/3483920

Chapter 1

General introduction

Both humans and songbirds are vocal learners that learn to produce their species-specific vocalizations early in life by exposure to the vocalizations of adult conspecifics (Doupe & Kuhl, 1999). For human speech and birdsong, better learning outcomes are often achieved with live, social, tutors than with audio-only exposure to vocalizations (speech: Bruner, 1983; Kuhl, Tsao, & Liu, 2003; Roseberry, Hirsh-Pasek, & Golinkoff, 2014, birdsong: reviewed in Baptista & Gaunt, 1997; Soma, 2011). Many researchers have argued that this is because social interactions between tutors and tutees are important in the vocal learning process (e.g. Beecher & Burt, 2004; Goldstein, King, & West, 2003; Kuhl, 2003, 2007). An open question, however, is whether and to what extent vocal learning from live tutors is also improved because live tutors enable tutees to both hear and see a tutor, instead of only hear a tutor (speech: Kuhl & Meltzoff 1982; Lewkowicz & Hansen-Tift 2012; Teinonen, Aslin, Alku, & Csibra 2008; Tenenbaum, Sobel, Sheinkopf, Malle, & Morgan 2015, birdsong: Beecher & Burt 2004; Derégnaucourt 2011; Slater, Eales, & Clayton 1988). Live tutoring, in other words, results in multimodal exposure to a tutor i.e. stimulation of multiple sensory modalities, while audio-only tutoring results in unimodal tutor exposure with stimulation of a single modality.

The simultaneous presentation of two stimuli in different modalities can improve signal perception compared to the presentation of one stimulus, as has been demonstrated in laboratory experiments in many taxonomic groups (reviewed in Rowe, 1999). Improved signal processing can occur when both stimuli are informative, but also when only one stimulus is relevant to the receiver, while the other is task-irrelevant and uninformative, but can draw the receiver's attention to the relevant stimulus (Alais, Newell, & Mamassian, 2010; Feenders, Kato, Borzeszkowski, & Klump, 2017; Rowe, 1999). The production of birdsong and speech are accompanied by sound-specific visual cues, such as songbirds' beak movements and human mouth movements. This makes speech and birdsong multimodal signals, i.e. signals that can be perceived through more than one sensory modality (Halfwerk et al., 2019; Higham & Hebets, 2013; Partan & Marler, 1999). Multi- compared to unimodal signalling can be beneficial for communication. For instance, if there is noise in one channel, information conveyed in the other channel can help receivers to identify the signal correctly (Partan & Marler, 2005). Additionally, multimodal signals are more likely to be detected by receivers than unimodal signals and receivers learn to recognize signals which contain multiple components (in one or multiple modalities) faster than single component signals (reviewed in Hebets & Papaj, 2005; Rowe, 1999). For example, big brown bats (*Eptesicus fuscus*) learn to avoid noxious fireflies faster with multi- than with unimodal warning signals (Leavell et al.,

2018). Exposure to the sound-specific visual cues accompanying vocalizations might have a faciliting effect on vocal learning, for instance by improving receivers' attention to the auditory signal or by directly facilitating learning of the underlying motor program to produce these vocalizations. The idea that audio-visual compared to audio-only exposure to vocalizing tutors could have a facilitating effect on vocal learning is supported by multiple lines of evidence from human and non-human animals.

In humans, visual information can affect speech perception in adults and infants. Visual exposure to a speaker's mouth and facial movements contributes to speech intelligibility, especially in noisy environments (Middelweerd & Plomp, 1987; Sumby & Pollack, 1954). Human infants of two months old already associate auditory and visually presented phonemes (Kuhl & Meltzoff, 1982). Besides, infants of around four months of age already perceive a different speech sound if auditory and visual speech cues are experimentally mismatched than they would perceive if the auditory and visual cue were presented separately (Burnham & Dodd, 2004). This is known as the McGurk effect and indicates that auditory and visual information are integrated into a multimodal percept (McGurk & Macdonald, 1976). Observational studies suggest that visual exposure to a speaking adult might play a role in early speech development. For instance, infants that fixate their gaze more on their mother's mouth during interaction at 6 months, show higher levels of expressive language (e.g. repeating sentences or naming objects) at age 2 (Young, Merin, Rogers, & Ozonoff, 2009). In addition, visual speech enhances learning of phoneme contrasts in 6-month-olds (Teinonen et al., 2008), and 12-month-olds pay more attention to a speaker's mouth when hearing a foreign language compared to their native language (Lewkowicz & Hansen-Tift, 2012). Infants that are born blind never experience visual exposure to speech. Although they acquire a speech system that seems comparable to that of sighted individuals, differences in the pronunciation of certain phonemes by blind and sighted individuals have been demonstrated (Ménard, Dupont, Baum, & Aubin, 2009). Moreover, for second language learning in adults, audio-visual training (with a speaker's mouth movements presented through videos of the speaker or through animation of a virtual head) improves the perception and production of unfamiliar speech contrasts more than audio-only training (e.g. Badin, Tarabalka, Elisei, & Bailly, 2010; Hazan, Sennema, Iba, & Faulkner, 2005; Hirata & Kelly, 2010; Liu, Massaro, Chen, Chan, & Perfetti, 2007; Wang, Hueber, & Badin, 2014). These studies suggest that it is worthwhile to experimentally investigate how early vocal development is affected by visual exposure to a tutor.

Like in humans, there are several studies in songbirds suggesting that visual stimulation might affect the perception and learning of vocalizations. Starlings (*Sturnus vulgaris*), for instance, show enhanced performance on a temporal order judgement task when visual cues were flanked by auditory stimuli (Feenders, Kato, Borzeszkowski, & Klump, 2017). This demonstrates that concurrent auditory and visual stimulation can influence stimulus perception. In nightingales (*Luscinia megarhynchos*), song presentation paired with stroboscope light flashes improved song learning compared to song presentation with no additional visual stimulation (Hultsch, Schleuss, & Todt, 1999). In the context of filial imprinting, young birds showed enhanced learning of an auditory stimulus when it was paired with a visual stimulus (van Kampen & Bolhuis, 1991; van Kampen & Bolhuis, 1993). These last two studies showed an effect of non-social and non-sound-specific visual stimulation on learning an auditory signal. It might be that any visual stimulation in addition to an auditory signal improves vocal learning of that signal equally, in which case visual exposure to a tutor would facilitate vocal learning to the same degree as non-social and non-sound-specific visual stimulation. It might also be that visual exposure to sound-specific movements has an additional facilitating effect on song learning, in which case seeing a singing tutor in addition to hearing song would facilitate song learning more than exposure to non-sound-specific visual stimulation.

Birdsong development provides a model system that can be used to experimentally investigate the effect of audio-visual compared to audio-only exposure to a tutor on vocal development. In this thesis, this question will be addressed investigating song development in zebra finches (*Taeniopygia guttata*), the primary experimental animal model for studies on vocal learning (Griffith & Buchanan, 2010; Mello, 2014). Zebra finch song production is accompanied by specific beak and body movements (Franz & Goller, 2002; Ohms, Snelderwaard, ten Cate, & Beckers, 2010; Ullrich, Norton, & Scharff, 2016; Williams, 2001). Individual zebra finches show stereotyped patterns of beak movements during song renditions (Goller, Mallinckrodt, & Torti, 2004; Williams, 2001). Changes in beak aperture are correlated with changes in song amplitude and frequency, and rapid changes in beak aperture occur mainly just before the onset of sound production and at rapid acoustic transitions during song (Goller et al., 2004; Ohms et al., 2010; Williams, 2001). A correlation between beak aperture and song frequency has been demonstrated in other songbird species as well (e.g. Podos, Southall, & Rossi-Santos, 2004; Westneat, Long, Hoese, & Nowicki, 1993). Zebra finches mainly combine singing with body movements as part of the courtship display performed in the presence of female conspe-

cifics (Ullrich et al., 2016; Williams, 2001). Audio-visual exposure to a singing tutor might affect song learning in zebra finches because it enables them to see these song-specific movements in addition to hearing the song.

The possibility that visual exposure to a singing tutor might (partially) explain improved song learning from live tutors has not been systematically studied yet. However, several observations suggest that young birds may attend to both auditory and visual information during song learning. In zebra finches, for instance, the beak movements of pupils show high similarity with those of their tutors compared to unfamiliar males (Williams, 2001), visual cues guide tutor choice (Mann & Slater, 1995; Mann, Slater, Eales, & Richards, 1991), and visual stimulation contingent with immature song production improves song learning in juvenile zebra finches (Carouso-Peck and Goldstein 2019). Moreover, zebra finches copy more song from a visible conspecific than from the playback of pre-recorded tutor song (Derégnaucourt, Poirier, van der Kant, & van der Linden, 2013) or from direct passive auditory exposure to a tutor through an opaque screen (Eales, 1989) or loudspeaker (Chen, Matheson, & Sakata, 2016). Although this suggests that seeing a tutor improves song learning, in these studies multimodal and social tutoring were confounded: in the tutoring treatments in which tutees could see their tutor, they could also visually interact with it. This makes it difficult to disentangle the effect of social and multimodal tutoring on zebra finch song learning. In this thesis, I therefore investigated the effect of multimodal tutoring on song learning success by using different tutoring methods where tutees could see the visual component of song production, without being able to visually interact with a tutor. To this end, I could utilise a robotic zebra finch that was jointly developed with collaborators from the Vrije Universiteit Amsterdam and that allowed standardized and controlled presentation of the auditory and visual component of song.

Thesis outline
To answer the question whether multi- compared to unimodal exposure to a tutor affects zebra finch song learning, this thesis first revisits the literature on zebra finch song learning experiments from the perspective of multi- versus unimodal tutoring and then describes three different song tutoring experiments.

In Chapter 2, the literature on zebra finch song learning under different tutoring treatments was reviewed to find out whether it supports the hypothesis that multi- compared to unimodal tutoring facilitates zebra finch song learning. Zebra finches copy more song from a live tutor than from auditory only expo-

sure to tutor song (Chen et al., 2016; Derégnaucourt et al., 2013; Eales, 1989). Several stimulus dimensions that differ between live and audio only tutoring have been experimentally tested for their effect on song learning, but it is as yet unclear what exactly the key facilitating factor of a live tutor is. The most favoured hypothesis for this difference is that a lack of social interaction with the tutor leads to poorer song copying from audio only playback than from a live tutor (Chen et al., 2016; Derégnaucourt et al., 2013; Slater, Eales, & Clayton, 1988). In this review, I investigated whether previous song learning studies have systematically controlled for multi- versus unimodal tutoring and whether their outcomes are in line with multi- compared to unimodal tutoring having an effect on the song learning process.

Chapter 3 describes a song tutoring experiment aimed at testing whether multi- compared to unimodal exposure to a live tutor facilitates zebra finch song learning. I investigated song learning in tutees that had visual exposure to an adult conspecific (the tutor) through a one-way mirror. These tutees thus had multimodal tutor exposure, but as the tutor could not see them, there was no visual social tutor-tutee interaction possible. I compared song learning in these tutees to that in tutees that did not have visual, but only auditory and therefore unimodal exposure to the tutor. I also investigated song learning in tutees that were raised in the same cage as the tutor, and that thus had multimodal tutor exposure, as well as the opportunity to visually and physically interact with the tutor. Tutees from all treatments could vocally interact with each other and the tutor and all tutees were housed with a female companion to avoid social isolation in the tutees that were not housed in the same cage as the tutor. If visual cues play a role in song learning, the tutees with multimodal tutor exposure would show improved tutor song copying compared to the tutees with unimodal tutor exposure. These results could be interpreted as support for the hypothesis that multimodal tutor exposure facilitates song learning, but an alternative, non-mutually exclusive, explanation could be that the tutor had facilitated song learning by providing visual feedback in response to the tutees' vocalizations. To prevent the possibility of the tutor providing visual feedback to tutees, I used artificial, instead of live, tutors in Chapter 4 and 5.

The studies described in Chapter 4 and 5 tested whether learning from passive, pre-recorded tutor song would be facilitated if tutees would at the same time be exposed to the visual cues accompanying the production of this song. In Chapter 4, these visual cues were presented through videos that had been adjusted for the zebra finch visual system using colour realistic imaginary and high speed video recordings and displays. I investigated song learning in tutees

that were exposed to a time aligned video of a tutor singing the song that they were at the same time auditorily exposed to. I compared this to learning in tutees that only heard this song and tutees that heard this song while they were exposed to the tutor video, but here the pixels were randomized and the frames were played in reversed order. The tutees that were presented to the original tutor video in addition to auditory song exposure were expected to show improved song learning compared to the other two tutoring conditions. While the realistic imaging techniques thus ensured a high fidelity audio-visual recording of a singing male, a video is only two-dimensional and lacking the depth of a real bird. This issue was addressed in the study described in Chapter 5.

In the study in Chapter 5, visual cues were presented by means of a three-dimensional robotic zebra finch producing beak and head movements time-aligned with the tutor song (RoboFinch, Simon et al., 2019). Tutees were exposed to the RoboFinch and their song learning was compared to that in two control groups: tutees exposed to the same tutor song without the robotic zebra finch present and tutees exposed to a robotic zebra finch that only started moving after auditory song presentation had finished. In this experiment, I also included a condition in which tutees were housed with a female companion while being exposed to song auditorily only, to find out whether the social isolation of the other tutees would negatively affect their song learning success. I expected the visual cues produced by the Robofinch and presented synchronized with the auditory song playback to facilitate song learning and to lead to a higher amount of tutor song copying than the other tutoring treatments.

Chapter 6 discusses the main conclusions with respect to the effect of audio-visual tutor exposure on song learning and discusses the results of this thesis in a broader perspective.

References

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory processing in review: From physiology to behaviour. Seeing and Perceiving (Vol. 23). https://doi.org/10.1163/187847510X488603

Badin, P., Tarabalka, Y., Elisei, F., & Bailly, G. (2010). Can you "read" tongue movements? Evaluation of the contribution of tongue display to speech understanding. Speech Communication, 52(6), 493–503. https://doi.org/10.1016/j.specom.2010.03.002

Baptista, L. F., & Gaunt, S. L. L. (1997). Social interaction and vocal development in birds. In C. T. Snowdon & M. Hausberger (Eds.), Social influences on vocal development (pp. 23–40). Cambridge, Cambridge University Press.

Beecher, M. D., & Burt, J. M. (2004). The role of social interaction in bird song

learning. Current Directions in Psychological Science, 13(6), 224–228. https://doi.org/10.1111/j.0963-7214.2004.00313.x

Bruner, J. (1983). Child's talk: Learning to use language. New York: W.W. Norton.

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. Developmental Psychobiology, 45(4), 204–220. https://doi.org/10.1002/dev.20032

Chen, Y., Matheson, L. E., & Sakata, J. T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. Proceedings of the National Academy of Sciences, 201522306. https://doi.org/10.1073/pnas.1522306113

Derégnaucourt, S. (2011). Birdsong learning in the laboratory, with especial reference to the song of the zebra finch (*Taeniopygia guttata*). Interaction Studies, 12, 324–350. https://doi.org/10.1075/is.12.2.07der

Derégnaucourt, S., Poirier, C., van der Kant, A., & van der Linden, A. (2013). Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. Journal of Physiology, 107, 210–218. https://doi.org/10.1016/j.jphysparis.2012.08.003

Doupe, A. J., & Kuhl, P. K. (1999). Bird song and human speech: common themes and mechanisms. Annu. Rev. Neurosci., 22, 567–631. https://doi.org/10.1146/annurev.neuro.22.1.567

Eales, L. A. (1987). Do zebra finch males that have been raised by another species still tend to select a conspecific song tutor? Animal Behaviour, 35(5), 1347–1355. https://doi.org/10.1016/S0003-3472(87)80007-6

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. Animal Behaviour, 37, 507–508. https://doi.org/10.1016/0003-3472(89)90097-3

Feenders, G., Kato, Y., Borzeszkowski, K. M., & Klump, G. M. (2017). Temporal ventriloquism effect in european starlings: evidence for two parallel processing pathways. Behavioral Neuroscience, 131(4), 337–347. https://doi.org/10.1037/bne0000200

Franz, M., & Goller, F. (2002). Respiratory units of motor production and song imitation in the zebra finch. Journal of Neurobiology, 51(2), 129–141. https://doi.org/10.1002/neu.10043

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. Proceedings of the National Academy of Sciences of the United States of America, 100(13), 8030–8035. https://doi.org/10.1073/pnas.1332441100

Goller, F., Mallinckrodt, M. J., & Torti, S. D. (2004). Beak gape dynamics, during song in the zebra finch. Journal of Neurobiology, 59(3), 289–303. https://doi.org/10.1002/neu.10327

Griffith, S. C., & Buchanan, K. L. (2010). The zebra finch : the ultimate Australian supermodel. Emu, 110, v–xii. https://doi.org/10.1071/MUv110n3_ED

Halfwerk, W., Varkevisser, J., Simon, R., Mendoza, E., Scharff, C., & Riebel, K. (2019). Toward testing for multimodal perception of mating signals. Frontiers in Ecology and Evolution, 7, 2013–2019. https://doi.org/10.3389/

fevo.2019.00124

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. Speech Communication, 47(3), 360–378. https://doi.org/10.1016/j.specom.2005.04.007

Higham, J. P., & Hebets, E. A. (2013). An introduction to multimodal communication. Behavioral Ecology and Sociobiology, 67(9), 1381–1388. https://doi.org/10.1007/s00265-013-1590-x

Hirata, Y., & Kelly, S. D. (2010). Effects of lips and hands on auditory learning of second-language speech sounds. Journal of Speech, Language, and Hearing Research : JSLHR, 53(April), 298–310. https://doi.org/10.1044/1092-4388(2009/08-0243)

Hultsch, H., Schleuss, F., & Todt, D. (1999). Auditory-visual stimulus pairing enhances perceptual learning in a songbird. Animal Behaviour, 58, 143–149. https://doi.org/10.1006/anbe.1999.1120

Kuhl, P. K. (2003). Human speech and birdsong: communication and the social brain. Proceedings of the National Academy of Sciences of the United States of America, 100(17), 9645–9646. https://doi.org/10.1073/pnas.1733998100

Kuhl, P. K. (2007). Is speech learning "gated" by the social brain? Developmental Science, 10(1), 110–120. https://doi.org/10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. Science, 218, 1138–1141. https://doi.org/10.1126/science.7146899

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. Proceedings of the National Academy of Sciences of the United States of America, 100(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. Proceedings of the National Academy of Sciences of the United States of America, 109(5), 1431–1436. https://doi.org/10.1073/pnas.1114783109

Liu, Y., Massaro, D. W., Chen, T. H., Chan, D., & Perfetti, C. (2007). Using visual speech for training chinese pronunciation: an in-vivo experiment. SLaTE Workshop on Speech and Language Technology in Education. ISCA Tutorial and Research Workshop. The Summit Inn, Farmington, Pennsylvania USA, (SLaTE), 29–32. Retrieved from http://www.isca-speech.org/archive_open/archive_papers/slate_2007/sle7_029.pdf

Mann, N. I., & Slater, P. J. B. (1995). Song tutor choice by zebra finches in aviaries. Animal Behaviour, 49(3), 811–820. https://doi.org/10.1016/0003-3472(95)80212-6

Mann, N. I., Slater, P. J. B., Eales, L. A., & Richards, C. (1991). The influence of visual stimuli on song tutor choice in the zebra finch, *Taeniopygia guttata*. Animal Behaviour, 42(2), 285–293. https://doi.org/10.1016/S0003-3472(05)80560-3

Mello, C. V. (2014). The zebra finch, *Taeniopygia guttata*: An avian model for inves-

tigating the neurobiological basis of vocal learning. Cold Spring Harbor Protocols, 2014(12), 1237–1242. https://doi.org/10.1101/pdb.emo084574

Ménard, L., Dupont, S., Baum, S. R., & Aubin, J. (2009). Production and perception of French vowels by congenitally blind adults and sighted adults. The Journal of the Acoustical Society of America, 126(3), 1406–1414. https://doi.org/10.1121/1.3158930

Middelweerd, M. J., & Plomp, R. (1987). The effect of speechreading on the speech-reception threshold of sentences in noise. Journal of the Acoustical Society of America, 82(6), 2145–2147. https://doi.org/10.1121/1.395659

Ohms, V. R., Snelderwaard, P. C., ten Cate, C., & Beckers, G. J. L. (2010). Vocal tract articulation in zebra finches. PLoS ONE, 5(7). https://doi.org/10.1371/ journal.pone.0011923

Partan, S., & Marler, P. (1999). Communication goes multimodal. Science, 283, 1272–1274. https://doi.org/0.1126/science.283.5406.1272

Podos, J., Southall, J. A., & Rossi-Santos, M. R. (2004). Vocal mechanics in Darwin's finches: Correlation of beak gape and song frequency. Journal of Experimental Biology, 207(4), 607–619. https://doi.org/10.1242/jeb.00770

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. Child Development, 85(3), 956–970. https://doi.org/10.1016/j.surg.2006.10.010.Use

Rowe, C. (1999). Receiver psychology and evolution of multicomponent signals. Animal Behaviour, 58, 921–931. https://doi.org/10.1006/anbe.1999.1242

Simon, R., Varkevisser, J., Mendoza, E., Hochradel, K., Scharff, C., Riebel, K., & Halfwerk, W. (2019). Development and application of a robotic zebra finch (RoboFinch) to study multimodal cues in vocal communication. PeerJ Preprints 7:E28004v3. https://doi.org/10.7287/peerj.preprints.28004v1

Slater, P. J. B., Eales, L. A., & Clayton, N. S. (1988). Song learning in zebra finches (*Taeniopygia guttata*): progress and prospects. Advances in the Study of Behaviour, 18, 1–34. https://doi.org/10.1016/S0065-3454(08)60308-3

Soma, M. F. (2011). Social factors in song learning: a review of Estrildid finch research. Ornithological Science, 10(2), 89–100. https://doi.org/10.2326/ osj.10.89

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. Journal of the Acoustical Society of America, 26(2), 212–215. https:// doi.org/10.1121/1.1907309

Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. Cognition, 108(3), 850–855. https://doi.org/10.1016/j.cognition.2008.05.009

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. Journal of Child Language, 42(06), 1173–1190. https://doi. org/10.1017/S0305000914000725

Ullrich, R., Norton, P., & Scharff, C. (2016). Waltzing Taeniopygia: integration of courtship song and dance in the domesticated Australian zebra finch. Animal

Behaviour, 112, 285–300. https://doi.org/10.1016/j.anbehav.2015.11.012

van Kampen, H. S., & Bolhuis, J. J. (1991). Auditory learning and filial imprinting in the chick. Behaviour, 117, 303–319. https://doi.org/10.1163/156853991X00607

van Kampen, H. S., & Bolhuis, J. J. (1993). Interaction between auditory and visual learning during filial imprinting. Animal Behaviour, 45, 623–625. https://doi.org/10.1006/anbe.1993.1074

Wang, X., Hueber, T., & Badin, P. (2014). On the use of an articulatory talking head for second language pronunciation training: the case of Chinese learners of French. 10th International Seminar on Speech Production, 449–452.

Westneat, M. W., Long, J. H., Hoese, W., & Nowicki, S. (1993). Kinematics of birdsong: functional correlation of cranial movements and acoustic features in sparrows. The Journal of Experimental Biology, 182, 147–171.

Williams, H. (2001). Choreography of song, dance and beak movements in the zebra finch (*Taeniopygia guttata*). The Journal of Experimental Biology, 204, 3497–3506.

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. Developmental Science, 12(5), 798–814. https://doi.org/10.1111/j.1467-7687.2009.00833.x