


**AUTHOR QUERY FORM**

	<p><b>Journal: YBRLN</b></p> <p><b>Article Number: 3855</b></p>	<p><b>Please e-mail or fax your responses and any corrections to:</b></p> <p><b>E-mail: <a href="mailto:corrections.essd@elsevier.sps.co.in">corrections.essd@elsevier.sps.co.in</a></b></p> <p><b>Fax: +31 2048 52799</b></p>
---	---	--

Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. To ensure fast publication of your paper please return your corrections within 48 hours.

For correction or revision of any artwork, please consult <http://www.elsevier.com/artworkinstructions>.

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Click on the 'Q' link to go to the location in the proof.

<b>Location in article</b>	<b>Query / Remark: <a href="#">click on the Q link to go</a> Please insert your reply or correction at the corresponding line in the proof</b>
<a href="#">Q1</a>	Please check the typeset of highlights is okay.
<a href="#">Q2</a>	References 'King et al. (2006), Steinsträter et al. (in preparation) and Moliadze et al. (2003)' are cited in the text but not listed. Kindly check.
<a href="#">Q3</a>	This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or, alternatively, delete it. Any reference not dealt with will be retained in this section.

Thank you for your assistance.

---

**Highlights**

---

► Statement of significance to the neuroscience of language: ► Extend the motor somatotopy of speech perception to the larynx  
Q1 **representation**. ► Functional dissociation between phonation and articulation also in **perception**. ► Demonstrate the role of the motor  
system also in vocal pitch **discrimination**.

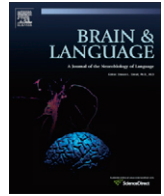
---



Contents lists available at ScienceDirect

# Brain & Language

journal homepage: [www.elsevier.com/locate/b&l](http://www.elsevier.com/locate/b&l)



## Vocal pitch discrimination in the motor system

Alessandro D'Ausilio<sup>a,b,\*</sup>, Ilaria Bufalari<sup>a</sup>, Paola Salmas<sup>b</sup>, Pierpaolo Busan<sup>a</sup>, Luciano Fadiga<sup>a,b,\*</sup>

<sup>a</sup>DSBTA, Section of Human Physiology, University of Ferrara, Ferrara, Italy

<sup>b</sup>IIT, The Italian Institute of Technology, Genova, Italy

### ARTICLE INFO

#### Article history:

Accepted 27 February 2011

Available online xxx

#### Keywords:

Motor cortex  
Motor theories of perception  
Speech comprehension  
TMS  
Motor resonance  
Larynx motor area  
Phonation  
Vocal pitch

### ABSTRACT

Speech production can be broadly separated into two distinct components: Phonation and Articulation. These two aspects require the efficient control of several phono-articulatory effectors. Speech is indeed generated by the vibration of the vocal-folds in the larynx (F0) followed by “filtering” by articulators, to select certain resonant frequencies out of that wave (F1, F2, F3, etc.). Recently it has been demonstrated that the motor representation of articulators (lips and tongue) participates in the discrimination of articulatory sounds (lips- and tongue-related speech sounds). Here we investigate whether the results obtained on articulatory sounds discrimination could be extended to phonation by applying a dual-pulse TMS protocol while subjects had to discriminate F0-shifted vocal utterances [a]. Stimulation over the larynx motor representation, compared to the control site (tongue/lips motor cortex), induced a reduction in RT for stimuli including a subtle pitch shift. We demonstrate that vocal pitch discrimination, in analogy with the articulatory component, requires the contribution of the motor system and that this effect is somatotopically organized.

© 2011 Published by Elsevier Inc.

### 1. Introduction

Sensori-motor integration requires specific brain circuits subserving coordinate transformation. The speech perception and production system is a particularly important and integrated instance of this kind. The intuition that articulatory goals may mediate perception was initially proposed long ago by the motor theory of speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) as well as the theory of analysis by synthesis (Stevens & Halle, 1967). The main reason for suggesting articulatory constraints for speech perception was the apparent lack of invariant cues in the acoustic signal to specify our phenomenal experience (Galantucci, Fowler, & Turvey, 2006). Generally speaking, both theories embrace a constructivist approach. Both of them indeed maintain that speech perception is mediated by constraints imposed by a sensori-motor model (via an inferential processes or internal simulation), mapping sensory input on the speech production system. A slightly different approach is the direct realist theory (Fowler, 1986). This theory proposes that, although there are no acoustic features that invariantly specify the units of speech, there are invariant properties in sensory stimulation that unambiguously specify the articulatory gestures, responsible for production, in a direct manner. This model, in fact, does not require any

inferential process. According to this approach what we perceive is not sensory in nature but directly relates to the articulatory gesture (Callan, Callan, Gamez, Sato, & Kawato, 2010).

Both the constructivist and the direct realist approach stress the role that phono-articulatory gestures might have for both production and perception of speech. Indeed a central theoretical tenet of both approaches is that speech classification is ultimately the recognition of the phono-articulatory gestures produced by the sender. At the same time several competing theories have emerged in the past decades suggesting that an exclusive sensory analysis is sufficient for classification (Diehl, Lotto, & Holt, 2004). Both views may enumerate a large number of evidences accumulated in a 50 years long debate regarding the supremacy of a sensori-motor or a purely sensory account. However, recent trends of research may now be added to the discussion by offering a renewed support for the former view.

In fact, on the computational side a recent and growing trend in automatic speech recognition literature acknowledge the beneficial role of articulatory features in improving phoneme/word classification (King et al., 2006). Furthermore, a simple prediction is the recruitment of the motor system during speech perception tasks. Many neuroimaging and neurophysiological studies have indeed showed that motor and premotor cortices become active while listening to speech sounds (Binder, Liebenthal, Possing, Medler, & Ward, 2004; Boatman & Miglioretti, 2005; Callan, Callan, Honda, & Masaki, 2000; Callan, Jones, Callan, & Akahane-Yamada, 2004; Callan et al., 2003, 2006, 2010; Londei et al., 2007, 2009; Möttönen, Jarvelainen, Sams, & Hari, 2004; Nishitani & Hari,

\* Corresponding authors at: DSBTA – Section of Human Physiology, University of Ferrara, Via Fossato di Mortara, 17/19 – 44100 Ferrara, Italy. Fax: +39 0532 455242.

E-mail addresses: [alessandro.dausilio@unife.it](mailto:alessandro.dausilio@unife.it) (A. D'Ausilio), [luciano.fadiga@unife.it](mailto:luciano.fadiga@unife.it) (L. Fadiga).

2002; Ojanen et al., 2005; Pulvermüller, Shtyrov, & Ilmoniemi, 2003; Pulvermüller et al., 2006; Shahin, Bishop, & Miller, 2009; Skipper, Goldin-Meadow, Nusbaum, & Small, 2007; Skipper, Nusbaum, & Small, 2005; Wang, Sereno, Jongman, & Hirsch, 2003; Wilson & Iacoboni, 2006; Wilson, Saygin, Sereno, & Iacoboni, 2004; and many others). Therefore, both computational and neuroimaging data seems to converge on the suggestion that motor computations/centers might take part in the speech classification process.

However, neuroimaging and most neurophysiological techniques use a correlational approach whereas transcranial Magnetic Stimulation (TMS) and direct cortical stimulation can directly demonstrate the activation of motor areas during perception of speech sounds. For instance, the discrimination of stop consonant voicing, place-of-articulation, syllable final vowel, steady-state vowels, pure tones, and frequency-modulated tones, is interfered by the direct electrical stimulation over superior temporal gyrus (Boatman, 2004). However, phoneme processing tasks such as phoneme monitoring, rhyming, and phoneme identification are interfered also by stimulation over the Inferior Frontal Gyrus. Interestingly, the inferior frontal gyrus might be recruited in syllable discrimination in patients with impaired comprehension under degraded listening conditions (Boatman & Miglioretti, 2005). On the other hand, studies employing TMS demonstrated that listening to tongue-produced speech sounds indeed activates the tongue motor representation (Fadiga, Craighero, Buccino, & Rizzolatti, 2002), whereas lips sounds activates the mouth motor area (Watkins, Strafella, & Paus, 2003). Furthermore TMS may be used to alter activity in motor centers and measure their causal involvement in perceptual tasks (D'Ausilio et al., 2009; Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007; Mättönen & Watkins, 2009; Sato, Tremblay, & Gracco, 2009). In a recent study we applied online focal TMS to the lips or tongue motor regions while subjects discriminated tongue- or lips-produced phonemes ([b] and [p] vs. [d] vs. [t]). TMS stimulation facilitated the discrimination of concordant phonemes with respect to the discordant items, thus showing that this effect is somatotopic and causally associated to the motor system (D'Ausilio et al., 2009). However, all these studies explored the motor somatotopy for speech perception focusing only on the articulatory component and neglecting the phonatory side of speech perception. This is mainly due to the fact that Liberman's theory derives from the observation that no invariant cues in the acoustic speech signal are present (i.e. coarticulation phenomena). However, this is not the case for vocal pitch processing on individual vowel, with pitch information clearly present in the acoustic signal. At the same time phonation is intimately linked to articulation at the production level.

In fact, phono-articulatory gestures include several articulators that might be controlled independently, with very different functions. Phonation is the control of laryngeal musculature in order to have the **vocal-folds** produce a quasi-periodic vibration (F0). Articulation instead, is the control of the tongue, lips, jaw, and other speech organs in order to filter resonant frequencies out of the F0 component. A speech signal, such as a vowel, is thus created by the combination of phonation and articulation. Typically, the larynx produce a fundamental frequency that is locally weighted by the resonances caused by the vocal tract. The F0 is the primary cue for pitch perception whereas subsequent formants (F1, F2, F3, etc.) characterize the vowel category (Ghazanfar & Rendall, 2008). Therefore, a strong associative link must exist between phonation, articulation and respiration, at the level of motor control. Here we seek to demonstrate whether such associative link translate into the involvement of larynx motor cortex in a vocal pitch discrimination task.

In this study we investigate the contribution of the motor system in the perception of phonated signals. Since laryngeal function

in speech production has a central importance specifically in determining vocal pitch, we investigated if the larynx representation of the motor cortex plays a role in vocal pitch discrimination. We used an approach similar to our previous TMS study on articulation (D'Ausilio et al., 2009), where we applied online TMS stimulation in order to experimentally alter activity in the motor system during a discrimination task. In the present study, subjects were required to discriminate whether pairs of vocal utterances [a] were the same or not. The two stimuli could parametrically differ in the F0 component height. We predict that if motor centers, and the larynx motor area specifically, have some role in discriminating vocal pitch, subjects' performance must be significantly different when TMS is applied.

## 2. Methods

### 2.1. Subjects

Ten healthy right-handed (measured with the Oldfield handedness questionnaire) subjects volunteered after giving informed consent and were compensated for their participation (mean age: 22.9, SD: 0.99; 7 **females**). None had any history of neurological, traumatic or psychiatric diseases and all of them had normal hearing. Procedures were approved by the local ethical committee in agreement with the Declaration of Helsinki.

### 2.2. Stimuli

Sound stimuli were recorded with a semi-professional microphone (AKG, C1000S) in a silent chamber and delivered to subjects via in-ear headphones. Stimuli were vocal recordings of a male actor producing a vowel sound [a]. The sound was cut and edited to last **600 ms** with a roughly constant **105 Hz** fundamental frequency. This vocal utterance was then pitch-shifted via Praat software (Version 5.1.05) by selectively moving the F0 formant. This procedure leaves unaltered all other formants and induces a perceptual pitch shift. F0 was initially shifted to  $\pm 3\%$ ,  $\pm 6\%$ ,  $\pm 9\%$  and  $\pm 12\%$ .

We then run a pilot experiment on **nine** subjects to select a subset of suitable pitch shifts. In fact, our aim was to avoid ceiling effects in behavioral recognition performance and to define intervals that were either very easy (close to 100%), very difficult (close to chance level) and an interval between them. Finally we had to consider that typical TMS experiments are relatively short and a limited number of magnetic pulses can be administered for safety reasons (Rossi et al., 2009). Therefore, we paired the original file with all other **pitch-shifted** samples and subjects had to decide whether the two differed or not. The  $\pm 3\%$  interval lead to an average of  $\sim 61\%$  of correct responses,  $\pm 6\%$  allowed  $\sim 85\%$ ,  $\pm 9\%$  enabled a very high performance of  $\sim 93\%$  and  $\pm 12\%$  with  $\sim 98\%$  was the easiest discrimination. We thus decided to use only the  $\pm 3\%$ ,  $\pm 6\%$  and  $\pm 9\%$  (7 samples in total, 1 original and 6 pitch-shifted) since the  $\pm 12\%$  contrast was too easy to discriminate and **could not** add much information to the experiment.

### 2.3. Task and procedure

Subjects were comfortably seated on a reclining armchair and the audio stimuli were presented via headphones. Subject' responses were acquired by a custom-made response pad and both stimuli presentation and data recording were controlled by an E-Prime (Psychology Software Tools, Inc.) script. The correct synchronization between auditory stimuli and TMS occurrence was preliminary tested by feeding both the PC sound-card output and the

TMS trigger to an external A/D board with an internal dedicated hardware clock (CED, micro1401).

Participants were first familiarized with the experimental stimuli and apparatus by completing 13 training trials (one each trial kind). Soon after that, we proceeded to the TMS mapping phase as described in Section 2.4. After completing these preparatory activities the experiment started. It consisted of two sessions (counterbalanced across subjects) differing only for the TMS stimulation site (see Section 2.4). Each session had 142 trials, half with TMS half without, randomly mixed. Each trial consisted of the presentation of two consecutive vocal utterance separated by 150 ms, that could either be the Same or Different. 70 trials contained the same audio stimuli and 72 a different pairs. The Same trials were constituted by 10 pairs each couple (7 samples in total, 1 original and 6 pitch-shifted) half with and half without TMS. Of those 72 Different trials, we had 24 trials for each pitch shift between the first and second stimuli ( $\pm 3\%$ ,  $\pm 6\%$ ,  $\pm 9\%$ ), half with and half without TMS (Fig. 1a). In Different trials, the first utterance was always the original one. Subjects' task was to decide, as fast and accurate as possible, whether the two utterances were the same or not by pressing one of the two buttons (the association between buttons and responses was balanced across subjects).

#### 2.4. TMS

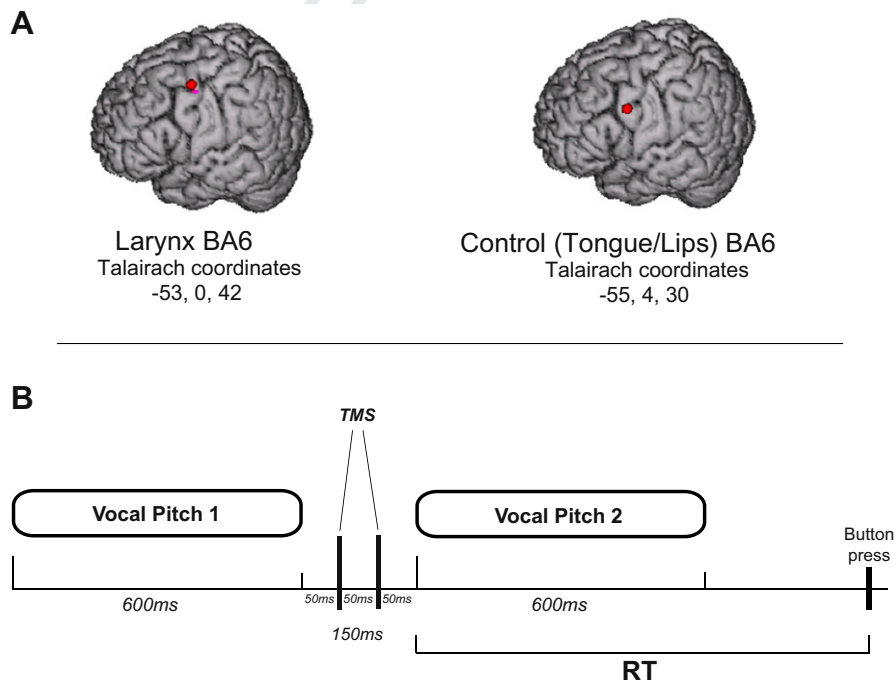
TMS stimulation was delivered through a figure-of-eight 25 mm coil and a Magstim Rapid stimulator (Magstim, Whitland, UK). The 25 mm coil was used to allow a more focal stimulation. Regarding spatial extent of TMS effects, we tried to be as specific as possible. In fact, we used an online protocol, instead of a more classical rTMS to avoid spreading of effects to adjacent areas. Motor Evoked Potentials (MEP) were recorded with a wireless EMG system (Aurion, ZeroWire EMG) by using a standard tendon-belly montage with Ag/AgCl electrodes. Signal was band-pass filtered (50–1000 Hz) and digitized (2 kHz). First Dorsal Interosseus (FDI) mapping and resting Motor Threshold (rMT) evaluation were assessed by using

standard protocols (5 out of 10 MEPs exceeding 50  $\mu$ V in peak-to-peak amplitude; Rossini et al., 1994).

Larynx and Control sites were, instead, localized by mixing a functional and a probabilistic method as in our previous TMS study (D'Ausilio et al., 2009). Specifically, for the Larynx premotor area we chose the Talairach coordinates revealed by a previous fMRI study (Larynx: -53, 0, 42; Brown, Ngan, & Liotti, 2008). The Control site corresponded to a premotor area lateral to the larynx and located on the overlap between lips and tongue motor representations (lips: -57, 4, 35; tongue: -53, 3, 24; Brown et al., 2008; Control stimulation site: -55, 4, 30; Fig. 1b). We choose this control site in order to stimulate speech effectors, unrelated to phonation and involved in articulation. In parallel, also FDI Talairach coordinates were taken from the literature (-37, -22, 54; Niyazov, Butler, Kadah, Epstein, & Hu, 2005).

In the following step, these coordinates (FDI, Larynx and Control) were transformed into the 10–20 EEG system space by using a popular web applet (Steinsträter et al., in preparation; web applet: <http://www.neuro03.uni-muenster.de/ger/t2tconv/conv3d.html>). Then, the distance between FDI/Larynx and FDI/Control was calculated. In each subject the FDI was functionally mapped, and then Larynx and Control were located according to these differential 10–20 EEG coordinates (Larynx: 9.2% of *nasion-inion* distance in the anterior direction and 5.2% of the inter-tragus distance in the lateral direction; Control: 9.1% anterior and 11.7% lateral; mean distance between Larynx and Control: 1.6 cm). Therefore, our localization method is based in part on a probabilistic approach, since we use normalized peak activations from previous studies. At the same time, these coordinates are scaled onto each subjects' individual external anatomy and also remapped according to individual functional location of the FDI muscle.

In the stimulated trials two pulses with 50 ms interval were delivered at 110% of the FDI rMT. This intensity was chosen to allow better spatial accuracy. In fact, the rMT for facial muscles has been reported to be higher than typical rMT for intrinsic hand muscles (Cruccu, Inghilleri, Berardelli, Romaniello, & Manfredi, 1997; Fadiga et al., 2002; Paradiso, Cunic, Gunraj, & Chen, 2005; Svens-



**Fig. 1.** Trial structure and stimulation sites. Panel A shows the stimulation sites for the experimental condition (Larynx pre-motor area) and the control area (tongue/lips pre-motor area) with relative Talairach coordinates. Panel B shows the trial structure, including auditory vocal stimuli and TMS timing.

son, Romaniello, Arendt-Nielsen, & Sessle, 2003; Svensson, Romaniello, Wang, Arendt-Nielsen, & Sessle, 2006; Watkins et al., 2003). According to these studies, the rMT for facial muscles is typically 10–20% higher than FDI rMT. The intensity was chosen to be just below the rMT for facial muscles (Crucchi et al., 1997; Paradiso et al., 2005; Svensson et al., 2003, 2006). Coil orientation was maintained at 45° with respect to the inter-hemispheric fissure, the handle pointing posteriorly. Pulses were given 50 ms and 100 ms after the end of the first stimulus; thus, the last TMS pulse occurred 50 ms prior to the second stimulus onset (see Fig. 1a).

2.5. Analysis

A 50% ratio between Same and Different trials was used to avoid response predictability. We were interested in subjects' ability to discriminate pitch differences, and RTs analyses were run on Different trials only. In fact, performance on Same trials was almost perfect, thus reaching a plateau level where we could not appreciate any change in RTs. All data was collapsed into three categories, ±3%, ±6% and ±9%, since in the pre-experiment there were no differences between positive and negative pitch shifts (+3% vs. -3%; +6% vs. -6%; +9% vs. -9%). RTs on correct trials were calculated from the beginning of second sound presentation. Subjects' performance was normalized by computing the mean RT ratio between TMS stimulated and un-stimulated trials. Here we similarly run a comparison between the two stimulation sites (Larynx, Control) in all three Conditions (±3%, ±6% and ±9%) with a series of two-tailed paired *t*-tests (corrected with the Bonferroni method). Effect-size was further computed with Cohen's *d* as well as Hedges' *g*.

For accuracy we conducted an analysis inspired by the Signal Detection Theory (SDT) by computing *d-prime* and beta ratio between TMS and no TMS, in both Same and Different trials. *d-Prime* values represent sensitivity of low level processing of speech cues whereas beta measures allow testing of possible response bias related to higher-level decision/categorization processes. For the accuracy data we compared the two sites (Larynx, Control) in all three Conditions (±3%, ±6% and ±9%) by means of a series of two-tailed paired *t*-tests (Bonferroni corrected for multiple comparisons). All statistical analyses were conducted by using Statistica 6.0 (StatSoft, Inc.) and MatLab (Mathworks, Inc.).

3. Results

Accuracy data did not yield any significant result (Table 2 contains raw accuracy data). The analysis on *d-primes* did not show any significant effect of stimulation site for all three conditions (Larynx: ±3% = 0.66 ± 0.25; ±6% = 0.76 ± 0.08; ±9% = 1.06 ± 0.09 - Control: ±3% = 1.09 ± 0.16; ±6% = 0.92 ± 0.10; ±9% = 0.97 ± 0.07 - *p*-values: ±3% = 0.25; ±6% = 0.1; ±9% = 0.47). Therefore, TMS application did not modify the overall discrimination sensitivity of subjects. Also beta values did not show significant results (Larynx: ±3% = -1.76 ± 1.66; ±6% = 3.98 ± 1.81; ±9% = -0.98 ± 2.24 - Control: ±3% = 1.87 ± 1.02; ±6% = -2.08 ± 1.61; ±9% = 0.97 ± 2.07 - *p*-values: ±3% = 0.17; ±6% = 0.1; ±9% = 0.51).

On the other hand, RTs data showed that TMS affected the discrimination speed of small pitch shifts when stimulating the Larynx pre-motor area (Fig. 2; Table 1 contains raw RT data). The discrimination of ±3% intervals led to a significant difference between the two stimulation sites, whereas ±6% and ±9% did not (±3%: *t*(9): -3.38, *p* = 0.008, Cohen's *d* = 1.31, Hedges' *g* = 1.72; ±6%: *t*(9): -1.04, *p* = 0.324, Cohen's *d* = 0.34, Hedges' *g* = 0.33; ±9%: *t*(9): -1.21, *p* = 0.256, Cohen's *d* = 0.55, Hedges' *g* = 0.56). This effect was in the direction of a behavioral facilitation since the TMS/NoTMS RT ratio was diminished by the application of TMS

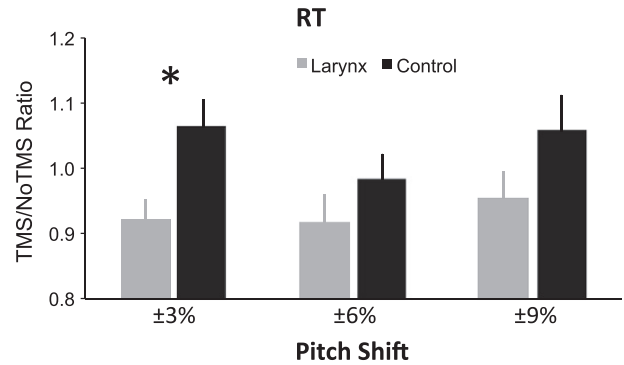


Fig. 2. RT Results. RTs data demonstrate that TMS applied over the Larynx pre-motor area affect subjects' discrimination of subtle vocal pitch shifts (±3%). RTs are shown in ratios between TMS and NoTMS trials. The asterisk denotes a statistically significant comparison and bars represent the standard error.

Table 1

RTs raw data. Reaction times data of the TMS experiment. Data includes trials with and without TMS, on both stimulation sites, as well as the standard error of the mean (SEM).

	Larynx area			Control area		
	±3%	±6%	±9%	±3%	±6%	±9%
TMS (ms)	436	403	387	508	430	398
NO-TMS (ms)	487	426	402	495	429	385
TMS (SEM)	43	57	51	45	54	51
NO-TMS (SEM)	47	46	44	44	45	43

Table 2

Accuracy raw data. Accuracy data for the TMS experiment including trials with and without TMS, on both stimulation sites, as well as the standard error of the mean (SEM).

	Larynx area			Control area		
	±3%	±6%	±9%	±3%	±6%	±9%
TMS (%)	70.83	83.33	92.50	72.50	97.50	98.33
NO-TMS (%)	68.33	90.83	90.83	68.33	93.33	96.67
TMS (SEM)	8.07	4.97	4.20	5.42	1.27	1.11
NO-TMS (SEM)	8.68	6.02	4.56	7.93	3.89	2.22

to the Larynx motor area (TMS/noTMS ratio at ±3%: Larynx = 0.89 ± 0.02 s.e.m., Control = 1.04 ± 0.03 s.e.m.; ±6%: Larynx = 0.93 ± 0.06 s.e.m.; Control = 0.98 ± 0.05 s.e.m.; ±9%: Larynx = 0.94 ± 0.04 s.e.m.; Control = 1.04 ± 0.07 s.e.m.). This means that the TMS applied to the larynx motor area reduced the RT when subjects had to discriminate small vocal pitch shifts.

4. Discussion

The idea that the motor system may play a role also in speech perception is not new and has been put forward several times from very different perspectives (See D'Ausilio, Craighero, & Fadiga, 2010). Classic aphasiology, for instance, very soon realized that a strict separation of production and perception abilities was rather artificial and far from clinical experience (Berker, Berker, & Smith, 1986). Several recent reports using different approaches seems to support an integrated and recurrent functional connection between anterior and posterior language areas, as well as the motor system, in a number of tasks ranging from simple phonetic discrimination to word comprehension (Pulvermüller & Fadiga, 2010).

The scope of the present study was to investigate low level processing of speech cues alternatively produced by different effectors

involved in speech production. Here we focused on the role played by the larynx motor representation in perceiving phonatory sounds. Interestingly, the localization of a larynx motor region has always been rather unclear (Brown et al., 2008). In fact, most studies were not differentiating phonation and articulation and in modern textbooks the larynx motor area is often located in the motor strip, ventral to the tongue representation, together with the pharynx (Penfield & Roberts, 1959). However, recent more accurate experiments demonstrated a different location, dorsal to the lips and ventral to the hand motor representation (Brown et al., 2009; Murphy et al., 1997; Olthoff, Baudewig, Kruse, & Dechent, 2008; Roedel et al., 2004).

Functionally speaking, most studies were not even differentiating phonation and respiration (Olthoff et al., 2008). However, respiration and vocal fold control are intimately linked and multiple representations may exist in the motor cortex for respiration (Ramsay et al., 1993). According to Ramsay et al. (1993), one was located in the trunk representation; the second focus was in the face area. The coordinates of the second site were quite proximate to the Larynx peaks reported in Brown et al. (2008) (Talairach: -44, -4, 40). These results raise indeed the possibility that there is a respiratory area very close to the larynx area, which might support respiratory/phonatory coordination during vocalization. According to this interpretation a clear-cut functional separation of respiration and phonation is difficult to observe at least with TMS stimulation in the pre-motor larynx area.

Nonetheless, beside the critical determination of where exactly the larynx motor area is located in the brain, it is worth investigating whether it plays also a role in perception (Brown et al., 2008) and if it functionally dissociates with respect to motor areas supporting articulatory sounds classification. In our previous TMS study we investigated the articulatory component of speech by discovering a double dissociation between tongue and lips motor area, when discriminating sounds produced mainly with the tongue or lips involvement (D'Ausilio et al., 2009). Here we analogously describe the contribution of the larynx motor area in the discrimination of subtle vocal pitch shift. The effect we show goes in the direction of performance facilitation, in line with our previous study (D'Ausilio et al., 2009). TMS-induced facilitation has been shown several times (Grosbras & Paus, 2003; Hayward, Goodwin, & Harmer, 2004; Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005; Töpper, Mottaghy, Brüggemann, Noth, & Huber, 1998) and might be explained by state-dependency effects of TMS (Silvanto, Muggleton, & Walsh, 2008), pulse timing (Moliadze et al., 2003) or stimulation protocol (Thut & Miniussi, 2009). In fact, TMS disrupts cortical processing for a limited time window, and heavily depends on stimulation protocol and activation state of the target area. However, animal models actually hold that, at rest, the effects of TMS turns into facilitation after a short period of time (50–100 ms) depending on stimulation intensity (Moliadze et al., 2003). In our study the second pulse was given 50msec prior to stimulus onset (and even a longer delay if we consider the putative motor cortex involvement in the task). The direction of our effects suggests that our TMS protocol is locally enhancing cortical activity.

However, differently for our previous study we found only an effect on RT but not on accuracy. The parallel effects on accuracy and RTs would have been a stronger proof in favor of a causal motor contribution to the task. However, we believe there's a clear reason for our present results. In fact, vocal pitch detection may require less contribution from the motor system (with respect to articulatory sounds analysis) since it may be "solved" via F0 detection on the auditory stream. In brief, the articulatory sound discrimination may effectively exploit the associated invariant motor features, whereas phonatory discrimination may predominantly take advantage from auditory invariants. Nevertheless, strong functional connection between sensori-motor representa-

tions of vocal pitch may exist, thus explaining our results on RTs only.

Therefore, our present results: (i) extend previous data on the motor somatotopy of speech perception to the larynx motor representation and (ii) hint to the functional dissociation between phonation and articulation also at the perceptual level. These two points are particularly relevant if we consider that the increasing number of studies, showing that the motor system might be involved in speech perception, typically do not consider the possible difference between articulation and phonation. For example, two recent studies, using offline repetitive TMS (rTMS), suggested that stimulating the ventral premotor cortex (vPM) impairs speech sound discrimination of syllables (Meister et al., 2007; Sato et al., 2009). However, the stimulation site of these two studies was borrowed from a previous fMRI study showing the overlap between the production and perception of syllables – with no distinction between phonation and articulation (Wilson et al., 2004). In fact, their stimulation site is very similar to the larynx motor area found by Brown and colleagues (Brown et al., 2008: -54, -2, 46; Wilson et al., 2004: -50, -6, 47; Meister et al., 2007: -53, -4, 49).

Moreover, we suggest that distinct motor representations may support the classification of speech signals. The behavioral evidence regarding the involvement of the larynx motor representation in the perception of specific speech cues nicely parallel the known dichotomy between phonation and articulation. In fact, when we turn our attention to speech production we realize that it is a complex phenomenon that requires the efficient control of several muscles and articulators, each one responsible for different aspects of speech sound spectrum. The larynx, in particular, is necessary for controlling vocal pitch, whereas other articulators (i.e. tongue, lips or jaws) are concerned with the filtering of the fundamental frequency, generated by the vocal folds (Ghazanfar & Rendall, 2008).

Phonation and articulation are indeed two separate facets of speech production/perception. In fact, these two components: (i) Are generated by different and independent effectors; (ii) Are controlled by distinct cortical motor representations; (iii) Play different, but complementary, functions in speech. This latter point, in our opinion, is the most critical and dense of implications for future research. As a matter of facts, vocal pitch modulations specifically carry a unique set of features extremely relevant at a behavioral level. Vocal pitch conveys important linguistic information such as the distinction between declarative and interrogative sentences or may help in revealing structures of ambiguous sentences. Moreover vocal pitch may also be used to decode speaker's gender, age and body size as well as critical supra-segmental information associated to emotional communication, such as emotional prosody or non-linguistic vocalizations (Wong, 2002). On one hand, gender, age and body size estimation may reveal important information on speaker identity, which can modulate social interactions. Pitch contour in linguistic or non-linguistic messages, on the other hand, can add important information on the emotional frame coloring communicative acts. Therefore, future investigations might indeed explore if and how the motor system contribute to the classification of the emotional contour surrounding verbal communication.

## 5. Uncited reference

Roy et al. (2008).

## Acknowledgment

L.F. is supported by Italian Ministry of Education; by the E.C. grants Contact, Robot-cub, Poeticon, Siempre and by Fondazione Cassa di Risparmio di Ferrara.

References

Berker, E. A., Berker, A. H., & Smith, A. (1986). Translation of Broca's 1865 report. Localization of speech in the third left frontal convolution. *Archives of Neurology*, 43, 1065–1072.

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, 7, 295–301.

Boatman, D. (2004). Cortical bases of speech perception: Evidence from functional lesion studies. *Cognition*, 92, 47–65.

Boatman, D. F., & Miglioretti, D. L. (2005). Cortical sites critical for speech discrimination in normal and impaired listeners. *Journal of Neuroscience*, 25, 5475–5480.

Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2009). The somatotopy of speech: Phonation and articulation in the human motor cortex. *Brain & Cognition*, 70, 31–41.

Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, 18, 837–845.

Callan, D., Callan, A., Gamez, M., Sato, M. A., & Kawato, M. (2010). Premotor cortex mediates perceptual performance. *Neuroimage*, 51(2), 844–858.

Callan, D., Callan, A., Honda, K., & Masaki, S. (2000). Single-sweep EEG analysis of neural processes underlying perception and production of vowels. *Cognitive Brain Research*, 10, 173–176.

Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *Neuroimage*, 22, 1182–1194.

Callan, D., Tajima, K., Callan, A., Kubo, R., Masaki, S., & Akahane-Yamada, R. (2003). Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *Neuroimage*, 19, 113–124.

Callan, D., Tsytarev, V., Hanakawa, T., Callan, A., Katsuhara, M., Fukuyama, H., et al. (2006). Song and speech: Brain regions involved with perception and covert production. *Neuroimage*, 31, 1327–1342.

Crucchi, G., Inghilleri, M., Berardelli, A., Romaniello, A., & Manfredi, M. (1997). Cortical mechanisms mediating the inhibitory period after magnetic stimulation of the facial motor area. *Muscle & Nerve*, 20, 418–424.

D'Ausilio, A., Craighero, L., & Fadiga, L. (2010). The contribution of the frontal lobe to the perception of speech. *Journal of Neurolinguistics*. doi:10.1016/j.jneuroling.2010.02.003.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, 19, 381–385.

Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, 55, 149–179.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 15, 399–402.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3–28.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin Review*, 13, 361–377.

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Current Biology*, 18, R457–R460.

Grosbras, M. H., & Paus, T. (2003). Transcranial magnetic stimulation of the human frontal eye field facilitates visual awareness. *European Journal of Neuroscience*, 18, 3121–3126.

Hayward, G., Goodwin, G. M., & Harmer, C. J. (2004). The role of the anterior cingulate cortex in the counting Stroop task. *Experimental Brain Research*, 154, 355–358.

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

Londei, A., D'Ausilio, A., Basso, D., Sestieri, C., Del Gratta, C., Romani, G. L., et al. (2007). Brain network for passive word listening as evaluated with ICA and Granger causality. *Brain Research Bulletin*, 72, 284–292.

Londei, A., D'Ausilio, A., Basso, D., Sestieri, C., Del Gratta, C., Romani, G. L., et al. (2009). Sensory-motor brain network connectivity for speech comprehension. *Human Brain Mapping*, 31, 567–580.

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, 17, 1692–1696.

Möttönen, R., Jarvelainen, J., Sams, M., & Hari, R. (2004). Viewing speech modulates activity in the left SI mouth cortex. *Neuroimage*, 24, 731–737.

Möttönen, R., & Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *Journal of Neuroscience*, 29, 9819–9825.

Murphy, K., Corfield, D. R., Guz, A., Fink, G. R., Wise, R. J., Harrison, J., et al. (1997). Cerebral areas associated with motor control of speech in humans. *Journal of Applied Physiology*, 83, 1438–1447.

Nishitani, N., & Hari, R. (2002). Viewing lip forms: Cortical dynamics. *Neuron*, 36, 1211–1220.

Niyazov, D. M., Butler, A. J., Kadah, Y. M., Epstein, C. M., & Hu, X. P. (2005). Functional magnetic resonance imaging and transcranial magnetic stimulation: Effects of

motor imagery, movement and coil orientation. *Clinical Neurophysiology*, 116, 1601–1610.

Ojanen, V., Mottonen, R., Pekkola, J., Jaaskelainen, I., Joensuu, R., Autti, T., et al. (2005). Processing of audiovisual speech in Broca's area. *Neuroimage*, 25(2), 333–338.

Olthoff, A., Baudewig, J., Kruse, E., & Dechent, P. (2008). Cortical sensorimotor control in vocalization: A functional magnetic resonance imaging study. *Laryngoscope*, 118(11), 2091–2096.

Paradiso, G. O., Cunic, D. I., Gunraj, C. A., & Chen, R. (2005). Representation of facial muscles in human motor cortex. *Journal of Physiology*, 567, 323–336.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11(5), 351–360.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, 21, 793–797.

Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 7865–7870.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2003). Spatiotemporal dynamics of neural language processing: An MEG study using minimum-norm current estimates. *Neuroimage*, 20, 1020–1025.

Ramsay, S. C., Adams, L., Murphy, K., Corfield, D. R., Grootoink, S., Bailey, D. L., et al. (1993). Regional cerebral blood flow during volitional expiration in man: A comparison with volitional inspiration. *Journal of Physiology*, 461, 85–101.

Roedel, R. M., Olthoff, A., Tergau, F., Simonyan, K., Kraemer, D., Markus, H., et al. (2004). Human cortical motor representation of the larynx as assessed by transcranial magnetic stimulation (TMS). *Laryngoscope*, 114, 918–922.

Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., & Safety of TMS Consensus Group (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, 120(12), 2008–2039.

Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., et al. (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: Basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalography Clinical Neurophysiology*, 91, 79–92.

Roy, A. C., Craighero, L., Fabbri-Destro, M., & Fadiga, L. (2008). Phonological and lexical motor facilitation during speech listening: A transcranial magnetic stimulation study. *Journal of Physiology – Paris*, 102, 101–105.

Sato, M., Tremblay, P., & Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain & Language*, 111, 1–7.

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage*, 44, 1133–1143.

Silvanto, J., Muggleton, N., & Walsh, V. (2008). State-dependency in brain stimulation studies of perception and cognition. *Trends in Cognitive Sciences*, 12, 447–454.

Skipper, J., Goldin-Meadow, S., Nusbaum, H., & Small, S. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain & Language*, 101, 260–277.

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

Stevens, K. N., & Halle, M. (1967). Remarks on analysis by synthesis and distinctive features. In W. Walthein-Dunn (Ed.), *Models for the perception of speech and visual form*. Cambridge, MA: MIT Press.

Svensson, P., Romaniello, A., Arendt-Nielsen, L., & Sessle, B. J. (2003). Plasticity in corticomotor control of the human tongue musculature induced by tongue-task training. *Experimental Brain Research*, 152, 42–51.

Svensson, P., Romaniello, A., Wang, K., Arendt-Nielsen, L., & Sessle, B. J. (2006). One hour of tongue-task training is associated with plasticity in corticomotor control of the human tongue musculature. *Experimental Brain Research*, 173, 165–173.

Thut, G., & Miniussi, C. (2009). New insights into rhythmic brain activity from TMS-EEG studies. *Trends in Cognitive Sciences*, 13, 182–189.

Töpper, R., Mottaghy, F. M., Brügmann, M., Noth, J., & Huber, W. (1998). Facilitation of picture naming by focal transcranial magnetic stimulation of Wernicke's area. *Experimental Brain Research*, 121, 371–378.

Wang, Y., Sereno, J., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of mandarin lexical tone. *Journal of Cognitive Neuroscience*, 15, 1019–1027.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41, 989–994.

Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage*, 33, 316–325.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7, 701–702.

Wong, P. C. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain Research Bulletin*, 59(2), 83–95.

574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657