

CLASSIFICATION: Social Sciences (Psychology)

ACTIVATION OF ARTICULATORY INFORMATION IN SPEECH PERCEPTION

Ivan Yuen,<sup>a</sup> Matthew H. Davis,<sup>b</sup> Marc Brysbaert,<sup>a,c</sup> & Kathleen Rastle<sup>a</sup>

a. Department of Psychology, Royal Holloway, University of London, Egham,

Surrey, TW20 0EX, United Kingdom

b. MRC Cognition and Brain Sciences Unit, Cambridge, CB2 7EF, United

Kingdom

c. Department of Psychology, Ghent University, 9000 Ghent, Belgium

MANUSCRIPT INFORMATION: 25 text pages, 2 figures, 3 tables

CORRESPONDING AUTHOR:

Kathleen Rastle

Department of Psychology

Royal Holloway, University of London

Egham, Surrey, TW20 0EX, UK

Phone: +44(0)1784 443 716

Fax: +44(0)1784 434 347

Email: [Kathy.Rastle@rhul.ac.uk](mailto:Kathy.Rastle@rhul.ac.uk)

## ABSTRACT

Emerging neurophysiological evidence indicates that motor systems are activated during the perception of speech, but whether this activity reflects basic processes underlying speech perception remains a matter of considerable debate. Our contribution to this debate is to report the first direct behavioral evidence that specific articulatory commands are activated *automatically* and *involuntarily* during speech perception. We used electropalatography to measure whether motor information activated from spoken distractors would yield specific distortions on the articulation of printed target syllables. Participants produced target syllables beginning with /k/ or /s/ while listening to the same syllables or to incongruent rhyming syllables beginning with /t/. Tongue-palate contact for target productions was measured during the articulatory closure of /k/ and during the frication of /s/. Results revealed ‘traces’ of the incongruent distractors on target productions, with the incongruent /t/-initial distractors inducing greater alveolar contact in the articulation of /k/ and /s/ than the congruent distractors. Two further experiments established that (a) the nature of this interference effect is dependent specifically on the articulatory properties of the spoken distractors; and (b) this interference effect is unique to spoken distractors and does not arise when distractors are presented in printed form. Results are discussed in terms of a broader emerging framework concerning the relationship between perception and action, whereby the perception of action entails activation of the motor system.

\body

One of the most exciting questions in the neuroscience of language concerns the involvement of the motor system in the perception of speech: is the motor system activated during speech perception and does it play a causal role? Key studies using fMRI have demonstrated that the brain regions involved in the perception of speech overlap with those involved in the production of speech<sup>1</sup> in a manner that appears to be articulator specific<sup>2</sup>. Similarly, studies using TMS have shown potentiation of motor cortex representations of the lip<sup>3</sup> and tongue<sup>4</sup> muscles when participants listen to speech. Finally, recent work using rTMS has revealed that disruption to regions of the premotor cortex impacts on perceptual discrimination of speech sounds<sup>5</sup> in a somatotopic manner<sup>6</sup>. These studies are all consistent with the proposal that the motor system is activated (and perhaps even essential) in the perception of speech.

However, while there is agreement that motor regions can be activated in speech perception studies, recent reviews of the literature have raised two important challenges over precisely what drives this activation<sup>7,8</sup>. The first challenge stems from the fact that neuroimaging data have been inconsistent, with relatively few studies showing motor activity at a whole-brain corrected level of significance compared to matched non-speech conditions<sup>7</sup>. Some investigators have attributed this inconsistency to the baselines used across studies, claiming that those studies using a complex acoustic baseline (e.g., musical rain, spectrally rotated speech) are less likely to yield motor activity unique to speech perception than those studies using a silent baseline<sup>7</sup>. The implication of this first challenge is that the motor activity sometimes observed in speech perception studies has nothing to do with the phonetic content of speech, but rather is driven by some acoustic event common to speech and non-speech sounds<sup>7,8</sup>. The second challenge concerns the possibility that motor activation arises

not as a result of basic processes underlying speech perception, but as a result of the requirements of certain tasks or listening situations<sup>7,8,9</sup>. The motor system could be recruited strategically to aid the interpretation of degraded speech, for example, or in conjunction with covert rehearsal processes used to improve performance on particular speech perception tasks. The implication of this second challenge is that there is a system for speech perception that is independent of the motor system and that evidence for motor involvement in speech perception arises as a consequence of neural processes that are outside of those normally and necessarily recruited in speech perception. This proposal is consistent with dual-pathway conceptions of speech perception embodied in dorsal–ventral stream accounts in which the primary mode of speech comprehension involves ventral auditory processes that are non-motoric<sup>10,11</sup>.

Our approach to this issue was to investigate the behavioral consequences of heard speech on the production of speech. Specifically, we hypothesized that if articulatory information is activated in speech perception, then this information should interfere with articulation in a scenario in which participants are asked to produce a target syllable while listening to a different auditory distractor. Rather than assessing whether conflicting auditory information introduces a delay in responding (a finding that would be difficult to attribute unambiguously to a particular level of processing) our approach was to investigate how an auditory distractor impacts upon the actual articulation of a different target. Our reasoning was that if articulatory information is activated in speech perception, then that information may interfere with speech production by introducing particular distortions of the target syllable that reflect the articulatory properties of the distractor.

This interference paradigm is characterized by two critical features that permit us to address the challenges described above in relation to neurophysiological

evidence for motor involvement in speech perception. First, the interference effects on speech production that we hypothesize are highly specific, reflecting particular phonetic properties of the spoken distractors. The observation of such specific distortions in the articulation of targets would be extremely difficult to reconcile with the view that motor activation in speech perception is driven by some acoustic event common to speech and non-speech sounds<sup>7</sup>. Second, the interference effects that we hypothesize constitute distortions of speech production, making the articulatory encoding of auditory distractors disadvantageous to performance. Thus, it would be difficult to argue that the motor system was being recruited strategically to assist participants with this particular speech perception task<sup>7,8,9</sup>. Indeed, if there is a system for speech perception that is independent of the motor system, then participants should be highly motivated to use it in this situation.

Recent studies in the area of experimental phonetics<sup>12,13</sup> provide evidence that laboratory-induced speech errors often consist of articulatory elements of the intended targets as well as the responses actually produced. This evidence suggests that the interference effects under examination here could also be manifest as articulatory blends of the target and auditory distractor. Thus, in order to quantify whether and how auditory distractors impact on the articulation of target syllables, we used a speech physiological technique known as electropalatography (EPG). EPG permits fine-grained analysis of spatio-temporal changes in the contact of the tongue against the roof of the mouth by sampling contact between the tongue and 62 sites on the palate every 10 ms (see Fig. 1, panel a). Thus, the use of EPG in this context will enable us to detect both transient and partial distortions of the articulatory gestures involved in speech production.

-- Figure 1 about here --

Targets and distractors were presented in the context of a specialized ‘tempo naming’ paradigm that enabled us to control the time taken for each response<sup>14</sup>. Participants were exposed to a series of five isochronous tones (500ms apart), on the fourth of which a printed target preceded immediately by an auditory distractor was presented. Participants were required to produce the printed target in line with the fifth tone. Following this sequence of events, participants made a phoneme monitoring judgment about the distractor (see Fig. 1, panel b). The tempo naming paradigm was used to force participants to pronounce the targets rapidly, thereby preventing them from resolving any interference caused by the distractors prior to the initiation of articulation<sup>14</sup>.

Participants produced target syllables beginning with /k/ or /s/ (e.g., /kib/) while listening to the same syllables (congruent distractors) or to rhyming syllables beginning with /t/ (incongruent distractors). Because the production of /t/ requires a complete alveolar closure formed with the tongue tip, our prediction was that the production of targets in the incongruent distractor condition would be characterized by greater tongue-palate contact in the first two rows of the palate (representing the alveolar region) than would the production of targets in the congruent distractor condition (see Fig. 1, panel c). The target phonemes /k/ and /s/ were chosen to establish the generality of this interference effect. Because the production of /k/ requires a complete velar closure formed with the tongue body, any effects of /t/-initial distractors would be manifest on a *different articulator* (i.e., with a different place of articulation) than that used to produce the target phoneme. In contrast, because the production of /s/ requires a narrow alveolar channel formed with the tongue tip, any effects of /t/-initial distractors would be manifest on the *same articulator* used to produce the target phoneme, suggesting that the effect of motor

activation can also be driven by dimensions such as the degree of constriction or manner of articulation. For both target phonemes, by analogy to models of manual action that propose a control system able to adjust motor programs in flight<sup>15</sup>, we predicted that the magnitude of this interference effect would be greater in the initial stages of target production than in the final stages of target production.

Two further experiments were conducted in order to rule out potential alternative explanations for any interference effects observed, and to home in on a theoretical account of those effects. The first additional experiment replaced the /t/-initial distractors with /g/-initial distractors for the /k/ targets and with /z/-initial distractors for the /s/ targets (these distractors being the voiced equivalents of their respective targets). This experiment thus allowed us to determine whether our interference effects could be attributed to the specific alveolar properties of the /t/-initial distractors rather than being a manifestation of phonetic incongruency between heard distractors and intended targets in general. The second additional experiment used printed /t/-initial distractors instead of auditory ones, thus allowing us to establish whether our interference effects were specific to the spoken modality, or also arise in another domain in which there are strong associative links to motor output. This experiment also allowed us to rule out covert rehearsal processes (possibly associated with short term memory demands of the phoneme monitoring judgments on auditory distractors) as an explanation for any interference effects observed. Though it is unlikely that these rehearsal processes would play a significant role in our interference paradigm (it has been known for at least 25 years that the concurrent articulation of a different target abolishes subvocal rehearsal<sup>16</sup>), this experiment provided a further check, since it is also well known that the modality of presentation makes no difference to whether these rehearsal processes are engaged<sup>16</sup>.

If such processes were associated with judgments on auditory distractors, then they should also be associated with judgments on printed distractors.

## RESULTS

### Experiment 1: Spoken Distraction

The prediction was that we would observe an effect of congruency, with incongruent /t/-initial distractors leaving articulatory traces of themselves on the production of /k/ and /s/ target phonemes, in the form of increased alveolar contact. Results (see Table 1 and Figure 2) confirmed this prediction in revealing a main effect of congruency on tongue contact in the first two rows of the palate,  $F(1,4)=32.76$ ,  $p=0.005$ . This main effect of congruency was modulated by an interaction with time,  $F(1,4)=8.89$ ,  $p=0.041$ , as the congruency effect was larger in the initial phase of target production than in the final phase. This congruency by time interaction was further modulated by target phoneme,  $F(1,4)=9.20$ ,  $p=.039$ , as the reduction in the size of the congruency effect toward the end of each phoneme was greater for /s/ targets than it was for /k/ targets, perhaps due to the greater duration of /s/ targets ( $M=145$  ms) than /k/ targets ( $M=80$  ms).<sup>\*</sup> Inspection of the individual subject data revealed that the congruency effect (and its predominance in the initial portion of targets) was apparent for each of the five participants (see Figure S1).

-- Table 1, Figure 2, and Figure S1 about here --

In order to establish that the articulatory consequences of the auditory distractors were highly specific (i.e., arising only in the first two rows of the palate, consistent with the production of /t/), we conducted exactly the same analysis using average tongue contact values obtained in Rows 3-8 of the palate. Because the tongue body (indicated by the region in Rows 3-8 of the palate) is not a primary articulator in

---

<sup>\*</sup> Further analyses revealed that the lexical status of targets and distractors did not influence the magnitude of the congruency effect in the initial portion of /k/ and /s/ target phonemes.



the production of /t/, we would not expect to observe any congruency effect in this analysis. Indeed, this analysis revealed no effect of congruency,  $F(1,4)=1.17$ ,  $p=.34$ , no interaction between congruency and time,  $F(1,4)=1.41$ ,  $p=.30$ , and no three-way interaction between congruency, time, and target phoneme,  $F(1,4)=.378$ ,  $p=.57$ . These analyses confirm that the articulatory traces left on target syllables by the incongruent /t/-initial distractors were highly specific: they were apparent in the alveolar region of the palate (Rows 1-2) but not in the other regions of the palate (Rows 3-8).

### Experiment 2: Spoken Non-Competing Distraction

Incongruent distractors in Experiment 2 began with /g/ (for /k/ targets) and /z/ (for /s/ targets). If the increased alveolar contact observed in the first experiment was the result of incongruity itself (instead of being a specific articulatory manifestation of the /t/-initial distractors), then the same congruency effect should be observed in this case. Results (see Table 2 and Figure 2) showed no effects of congruency, or interactions between congruency, time, and target phoneme on tongue contact in the first two rows of the palate (all  $F_s < 1$ ). Further, cross-experiment statistical comparisons revealed a larger effect of congruency in the first experiment than in this experiment,  $F(1,4)=12.74$ ,  $p=.023$ . These data confirm that the increased alveolar contact observed in Experiment 1 was due specifically to the presence of an incongruent /t/-initial distractor characterized by a complete alveolar closure.

-- Table 2 about here --

### Experiment 3: Printed Distraction

This experiment was conducted (a) to test whether the interference effects observed in Experiment 1 also arise in a domain in which there are strong associative links to the motor system; and (b) to rule out subvocal rehearsal of distractors as an

explanation for the interference effects observed. Results (see Table 3 and Figure 2) showed no effects of congruency ( $F < 1$ ), and no interactions between congruency and time,  $F(1,4) = 1.96$ ,  $p = .23$ , or between congruency, time, and target phoneme,  $F(1,4) = 2.57$ ,  $p = .18$  on tongue contact in the first two rows of the palate. Further, cross-experiment statistical comparisons revealed a larger congruency effect in the first experiment than in this experiment,  $F(1,4) = 10.34$ ,  $p = .032$ . These results confirm that the interference effects observed in Experiment 1 are unique to the perception of spoken distractors and undermine any account of those effects based on covert rehearsal.

-- Table 3 about here --

## DISCUSSION

The research reported in this article contributes to debate regarding the involvement of motor systems in the perception of speech. Though recent neurophysiological findings have demonstrated that motor and premotor regions can be activated in speech perception<sup>1,2,3,4,5,6</sup>, many have argued strongly that this activation arises as a result of neural processes outside the realm of normal speech perception<sup>7,8,9</sup>. These include processes associated with the interpretation of complex acoustic signals not specific to speech<sup>7</sup> or processes associated with the strategic demands of particular speech perception tasks<sup>7,8,9</sup>.

Our approach to this problem was to hypothesize that if articulatory information is activated in speech perception, then that information should interfere with the articulation of a different target in speech production. Participants were asked to read aloud printed targets beginning with /k/ and /s/ under deadline conditions, while also listening to an auditory distractor that matched the target (congruent) or that rhymed with the target but began with /t/ (incongruent). Specialized artificial

palates were custom-fitted to each participant so that we could monitor tongue-palate contact during articulation of the targets. Results showed that articulation of the target phonemes /k/ and /s/ was modified by the phonetic characteristics of the auditory distractors. Specifically, the articulation of both /k/ and /s/ phonemes revealed significantly greater tongue-palate contact in the alveolar region of the palate when distractors began with /t/ (a phoneme characterized by a complete alveolar closure) than when distractors were identical to the targets. This interference effect was particularly pronounced in the initial portions of each target segment. Further experiments established that this interference effect was caused by the specific articulatory characteristics of the /t/-initial distractors, and that it reflects processes unique to the perception of spoken (as opposed to printed) distractors.

The interference effects that we have observed permit us to address the two challenges described in the introduction regarding the nature of motor activation in speech perception. In respect of the first challenge, the fact that these interference effects were highly specific (reflecting the articulatory properties of the spoken distractors) lends substantial weight to the argument that motor activation in speech perception is driven by the phonetic content of speech, rather than some acoustic event common to speech and non-speech sounds<sup>7</sup>. In respect of the second challenge, what is critical is that we observed these effects in a situation in which the articulatory encoding of auditory distractors was disadvantageous to performance (because it resulted in distorted speech tokens). Participants could not inhibit the articulatory encoding of the auditory distractors, implying that articulatory information is activated *automatically* and *involuntarily* in speech perception, rather than being recruited strategically to assist participants with challenging listening situations or particular tasks<sup>7,8</sup>. Though we cannot rule out dual-pathway conceptions of speech

perception embodied in dorsal–ventral stream accounts (in which the primary route for speech comprehension involves ventral auditory processes that are non-motoric),<sup>10,11</sup> our evidence would suggest that the operation of the dorsal pathway linking auditory regions to prefrontal and motor cortex cannot be suppressed even under task conditions that would directly favor this.

Our interpretation of the interference effects reported in this article is that motor programs were activated by the auditory distractors, and when these conflicted with the motor programs activated by the printed targets, the motor programs combined resulting in an intermediate articulatory outcome. The fact that these interference effects were unique to the spoken modality demonstrates that speech has a privileged status insofar as the mapping to articulation is concerned. However, while our work demonstrates the tight coupling between auditory and articulatory information, it does not allow us to conclude that motor processes are *required* in speech perception (as in single pathway architectures in which motor representations mediate the acoustic and linguistic processing of speech<sup>17,18</sup>). Such inferences about causality can be drawn conclusively only by studying speech perception in situations in which motor processes are impaired through temporary lesions (as in TMS<sup>5,6</sup>) or as a result of brain injury<sup>19,20</sup>. However, further work establishing the temporal character of motor activation in speech perception could also be important in this respect. If motor activation were a necessary precursor to speech comprehension, it would need to occur very rapidly indeed given existing data concerning the time-course of spoken word recognition<sup>21</sup>. This information could be gleaned in further work using the interference paradigm, for example, by varying the interval between distractor and target and measuring the resulting influence on the magnitude of interference.

Some may claim that evidence for motor involvement in speech perception is unsurprising given the strong co-occurrence between heard speech and produced speech<sup>7</sup>. Specifically, it has been suggested that functional connections between neuron groups involved in articulatory and acoustic processing emerge simply as a result of associative learning<sup>22</sup> in the same way as distributed neuronal assemblies develop to bind the semantic representations of action words (e.g., kick, lick) and the motor representations used for implementing those actions<sup>23</sup>. However, if the acoustic-to-articulatory links implicated by our findings emerged as a result of associative learning, then surely we should have observed articulatory interference effects when distractors were presented in printed form. Indeed, it is well accepted that there is a strong association between text and speech – a relationship that is central to adult skilled reading<sup>24</sup> and reading development<sup>25</sup>. However, despite this strong association (and unlike the situation in which distractors were presented in the auditory modality) there is no evidence that articulatory information was activated when distractors were presented in this manner.

Though our data suggest that the link between speech perception and motor gestures is not one that can be explained by simple associative learning processes, we would not like to claim that this link arises through a specialized linguistic module encapsulated from other perceptual processes<sup>26</sup>. Rather, we prefer to interpret our data in the context of a broader emerging framework, whereby the perception of action entails activation of the motor system. It has not escaped our attention that effects similar to the ones that we observed in speech have already been observed in research using kinematic analyses to investigate reaching and grasping behavior<sup>27,28</sup>. In a manner somewhat analogous to our findings, one study showed that grip apertures for grasping an apple were smaller when a cherry was presented as a

distractor fruit than when an orange was presented as a distractor fruit<sup>27</sup>. It seems unlikely to us that these effects should be interpreted as arising from wholly different mechanisms to those that we observed.

The finding that the articulatory characteristics of auditory distractors exert a highly-specific interference effect on the articulation of target syllables argues strongly for the automatic recruitment of motor systems in speech perception yet also raises numerous possibilities for future research. One interesting question concerns whether articulatory information is extracted from synthesized speech. On the basis of recent behavioral<sup>29</sup> and neurophysiological<sup>30</sup> evidence suggesting that the mirror system underlying the perception of manual action is biologically tuned, it seems possible that the interference effects on articulation reported here would be observed only when auditory distractors comprise human speech. Having established a behavioral diagnostic for the activation of articulatory information in speech perception, we are now in a position to answer this and other important questions.

## MATERIALS AND METHODS

### Participants.

Four female and one male between the ages of 21 and 50 participated in the experiments (sample sizes of five or fewer participants are typical of EPG experiments<sup>31,32</sup>). Participants were all native speakers of English, had normal or corrected-to-normal vision, and had no known language, speech, or hearing deficits. Participants all signed a consent form approved by the Royal Holloway, University of London ethics committee.

### Materials.

Twenty-four target syllables with a consonant-vowel and consonant-vowel-consonant structure were created. These syllables reflected every combination of two

onsets (/k/ and /s/), three vowels (/i/, /a/, and /u/), and four codas (/p/, /b/, /m/, and null). During the experiment, participants were instructed to produce these syllables following a schwa (e.g., /ə kup/), and they appeared visually as the second constituent of a two-syllable stimulus (e.g., a koop).

Each of the target syllables was paired with a congruent distractor and an incongruent distractor. The congruent distractors were always phonologically identical to the targets. The nature of the incongruent distractors varied as a function of experiment: in Experiments 1 and 3 they rhymed with the targets but began with /t/; in Experiment 2 they rhymed with the targets but began with /g/ (in the case of /k/ targets) and /z/ (in the case of /s/ targets). Auditory distractors were produced by a female speaker of Standard Southern British English, and were recorded directly to the hard drive of a PC at a sampling rate of 22 KHz. They were approximately 400 ms in duration.

Each target syllable was presented twice with its congruent and incongruent distractors, yielding a total of 96 productions per participant per experiment. Stimuli were presented in a different random order for each participant.

#### Apparatus.

Participants were tested in a sound-treated room in the Department of Psychology, Royal Holloway University of London using a PC-based WinEPG electropalatography system (Articulate Instruments).

Each participant was fitted with an artificial acrylic palate suitable for use with the WinEPG system. The artificial palates were made from plaster cast impressions of each participant's mouth. Embedded in each palate were 62 electrodes in eight evenly-spaced rows from the front to the back of the palate (with six electrodes in the front row, and eight electrodes in every other row). These electrodes registered

tongue-palate contact in a binary manner, from behind the upper front teeth to the junction of the hard and soft palate. The wires connected to each electrode were bundled into two thin tubes that were fed out from the corners of the mouth and connected to the control box of the WinEPG system. Participants had practiced using their palates while speaking, and were subject to a 30 minute acclimatization exercise prior to the testing session.

Stimulus presentation was controlled by the DMDX software<sup>33</sup> running on a Pentium 4 PC. The recording of acoustic and EPG data was controlled by the Articulate Assistant software (Version 1.12) running on a separate Pentium 4 PC at sampling rates of 22 KHz and 100 Hz, respectively. Acoustic data were collected using an AKG microphone placed approximately 10-12 cm diagonally from the center of the lower lip. Stimulus presentation was integrated with the acoustic and EPG data collection, such that each auditory stimulus delivered to participants was also delivered to the acoustic record.

### Procedure.

Participants were seated approximately 16 inches (about 40cm) from the computer monitor and were asked to produce the syllables /di/ and /da/ several times each while the EPG apparatus was calibrated. Participants were given two practice blocks prior to each experiment.

Each experimental trial comprised a series of five 50 ms tones separated by intervals of 500 ms. On the fourth tone, a target stimulus appeared on screen. The appearance of this target stimulus was immediately preceded by an auditory or printed distractor (printed in a different color to the target). Participants were instructed to produce the target stimulus (including the initial schwa) on the fifth tone. To ensure that participants attended to the distractors, they were required to make a judgment



about the distractor following their spoken response, in which they were asked about the presence of a particular sound or letter. These judgments required a ‘YES’ response on 50% of trials and a ‘NO’ response on the other 50% of trials.

#### Data Preparation.

Targets were scrutinized for accuracy, with tokens excluded from the analyses based on perceptual criteria (e.g., producing the wrong onset, vowel, or coda). Four tokens were removed from Experiment 1 (0.83% of the data), eight tokens were removed from Experiment 2 (1.67% of the data), and seven tokens were removed from Experiment 3 (1.46% of the data).

Target utterances were annotated using waveform and spectrographic data. The initial segment of the /s/ targets, defined by a period of acoustic frication resulting from a turbulent airstream being forced through a narrow channel, was labeled using various acoustic markers. The onset of /s/ was defined by the beginning of high-frequency noise in the spectrogram, while the offset of /s/ was defined by (a) the end of high-frequency noise in the spectrogram; and (b) the onset of the glottal pulse and first formant of the following vowel in the spectrogram. For the /k/ targets, we were interested in the closure phase of production, during which time the vocal tract is occluded completely and there is no acoustic radiation from the lips. Because the target syllables were preceded by a schwa, the onset of the /k/ closure was defined by the minimal amplitude in the waveform and the onset of silence in the spectrogram. The offset of the closure period was defined by the sudden burst associated with the release of the articulatory closure. This burst was visible on the spectrogram as the onset of mid-high frequency energy and on the waveform as an increase in amplitude.

The onset and offset boundaries of /s/ and /k/ were used to extract these target segments from each utterance. The duration of each segment was then used to calculate two equally-spaced time points within each segment, representing one-third and two-thirds of the duration of the segment. In order to test our hypothesis that auditory /t/-initial distractors would leave alveolar traces of themselves on the articulation of target syllables, tongue contact values in Rows 1 and 2 of the palate (the rows immediately behind the upper front teeth) were extracted at each of four time points (onset, 1/3, 2/3, offset) for each of the /k/ and /s/ target segments for each of the experiments. Palates for the first two time points were averaged and treated as the ‘initial’ portion of the segment, while palates for the last two time points were averaged and treated as the ‘final’ portion of the segment. Similar values were obtained for tongue contact in Rows 3 through 8 of the palate, so that the specificity of any effect of congruency in the alveolar region of the palate could be established.

#### Statistical Analysis.

EPG data were analyzed in a repeated-measures analysis of variance (ANOVA) with target phoneme (/k/ or /s/), congruency (congruent or incongruent), and time (initial or final portion of target phonemes) as factors. Because we were interested specifically in the congruency effect, only main effects and interactions involving that factor are reported.

## ACKNOWLEDGEMENTS

This work was supported by the Biotechnology and Biological Sciences Research Council (BB/E003419/1) and the Medical Research Council (U.1055.04.013.1.1). Correspondence concerning this article may be addressed to Kathleen Rastle, Department of Psychology, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK, email: [Kathy.Rastle@rhul.ac.uk](mailto:Kathy.Rastle@rhul.ac.uk).

## REFERENCES

1. Wilson, SM, Saygin, AP, Sereno, MI, & Iacoboni, M (2004). Listening to speech activates motor areas involved in speech production. *Nat Neurosci* 7: 701-702.
2. Pulvermüller, F, Huss, M, Kherif, F, Moscoso del Prado Martin, F, Hauk, O, & Shtyrov, Y (2006). Motor cortex maps articulatory features of speech sounds. *P Natl Acad Sci USA* 103: 7865-7870.
3. Watkins, KE, Strafella, AP, & Paus, T (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41: 989-994.
4. Fadiga, L, Craighero, L, Buccino, G, & Rizzolatti, G (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur J Neurosci* 15: 399-402.
5. Meister, IG, Wilson, SM, Deblieck, C, Wu, AD, & Iacoboni, M (2007). The essential role of premotor cortex in speech perception. *Curr Biol* 17: 1692-1696.
6. D'Ausilio, A, Pulvermüller, F, Salmas, P, Bufalari, I, Begliomini, C, & Fadiga, L (2009). The motor somatotopy of speech perception. *Curr Biol* 19: 1-5.
7. Scott, SK, McGettigan, C, & Eisner, F (2009). A little more conversation, a little less action – candidate roles for the motor cortex in speech perception. *Nat Rev Neurosci* 10: 295-302.
8. Lotto, AJ, Hickok, GS, & Holt, LL (2009). Reflections on mirror neurons and speech perception. *Trends Cogn Sci* 13: 110-114.
9. Hickok, G (2009). Speech perception does not rely on motor cortex: Response to D'Ausilio et al. *Commentary available on* [http://www.cell.com/current-biology/comments\\_Dausilio](http://www.cell.com/current-biology/comments_Dausilio).

10. Hickok, G, & Poeppel, D (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92: 67-99.
11. Scott, SK & Johnsrude, IS (2003). The neuroanatomical and functional organization of speech perception. *Trends Neurosci* 26: 100-107.
12. Goldrick, M, & Blumstein, SE (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Lang Cognitive Proc* 21: 649-683.
13. Pouplier, M (2007). Tongue kinematics during utterances elicited with the SLIP technique. *Lang Speech* 50: 311-341.
14. Kello, CT & Plaut, DC (2000). Strategic control in word reading: Evidence from speeded responding in the tempo naming task. *J Exp Psychol Learn* 26: 719-750.
15. Glover, S (2004). Separate visual representations in the planning and control of action. *Behav Brain Sci* 27: 3-24.
16. Baddeley, AD, Lewis, VJ, & Vallar, G (1984). Exploring the articulatory loop. *Q J Exp Psychol* 36: 233-252.
17. Liberman, AM & Whalen, DH (2000). On the relation of speech to language. *Trends Cogn Sci* 4: 187-196.
18. Hickok, G, Holt, LL, & Lotto, AJ (2009). Response to Wilson: What does motor cortex contribute to speech perception? *Trends Cogn Sci* 13: 330-331.
19. Moineau, S, Dronkers, NF, & Bates, E (2005). Exploring the processing continuum of single-word comprehension in aphasia. *J Speech Lang Hear R*, 48: 884-896.
20. Utman, JA, Blumstein, SE, & Sullivan, K (2001). Mapping from sound to meaning: Reduced lexical activation in Broca's aphasics. *Brain Lang*, 79: 444-472.

21. Marslen-Wilson, WD & Tyler, LK (1980). The temporal structure of spoken language understanding. *Cognition* 8: 1-71.
22. Pulvermüller, F (2008). *The neuroscience of language: On brain circuits of words and serial order*. Cambridge: Cambridge University Press.
23. Pulvermüller, F (2005). Brain mechanisms linking language and action. *Nat Rev Neurosci* 6: 576-582.
24. Rastle, K & Brysbaert, M (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychol* 53: 97-145.
25. Rack, JP, Snowling, MJ, & Olson, RK (1992). The nonword reading deficit in developmental dyslexia: A review. *Read Res Quart* 27: 29-53.
26. Liberman, AM & Mattingly, IG (1985). The motor theory of speech perception revised. *Cognition* 21: 1-36.
27. Castiello, U (1996). Grasping a fruit: Selection for action. *J Exp Psychol Human* 22: 582-603.
28. Glover S & Dixon P (2002). Semantics affect the planning but not control of grasping. *Exp Brain Res* 146: 383-387.
29. Kilner, JM, Paulignan, Y, & Blakemore, SJ, (2003). An interference effect of observed biological movement on action. *Curr Biol* 13: 522-525.
30. Tai, YF, Scherfler, C, Brooks, DJ, Sawamoto, N, & Castiello, U (2004). The human premotor cortex is a 'mirror' only for biological actions. *Curr Biol* 14, 117-120.
31. Narayanan, SS, Alwan, AA, & Haker, K (1997). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. 1. The laterals. *J Acoust Soc Am* 101: 1064-1077.

32. Rastle, K., Harrington, JC, Coltheart, M., & Palethorpe, S. (2000). Reading aloud begins when the computation of phonology is complete. *J Exp Psychol Human* 26: 1178-1191.
33. Forster, KI & Forster, JC (2003). DMDX: A windows display program with millisecond accuracy. *Behav Res Meth Ins C* 35: 116-124.
34. Loftus, GR & Masson, MEJ (1994). Using confidence intervals in within-subject designs. *Psychon B Rev* 1, 476-490.

## FIGURE LEGENDS

Figure 1. Experimental predictions and design. (a) Sample palate worn by participants with alveolar electrodes highlighted; (b) Structure of an experimental trial; (c) Simplified schematic of the experimental predictions.

Figure 2. The congruency effect in each experiment as a function of time. This graph shows mean proportion tongue contact on alveolar electrodes during the initial and final portions of annotated segments collapsed across target phoneme. Error bars show the standard error of the mean for each experiment after removing between-subject variance (suitable for repeated measures comparisons<sup>34</sup>).

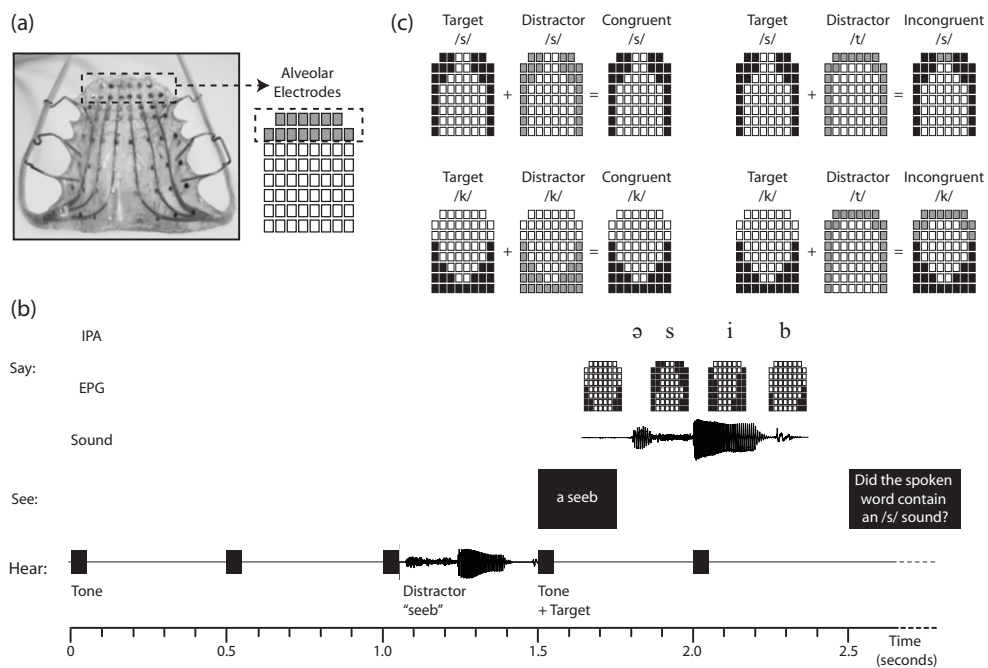


## TABLE LEGENDS

Table 1. Tongue contact values for each target phoneme as a function of congruency and time for Experiment 1. Expressed as the number of active contacts in the first two rows of the palate divided by the total number of contacts in these rows (N=14).

Table 2. Tongue contact values for each target phoneme as a function of congruency and time for Experiment 2. Expressed as the number of active contacts in the first two rows of the palate divided by the total number of contacts in these rows (N=14).

Table 3. Tongue contact values for each target phoneme as a function of congruency and time for Experiment 3. Expressed as the number of active contacts in the first two rows of the palate divided by the total number of contacts in these rows (N=14).



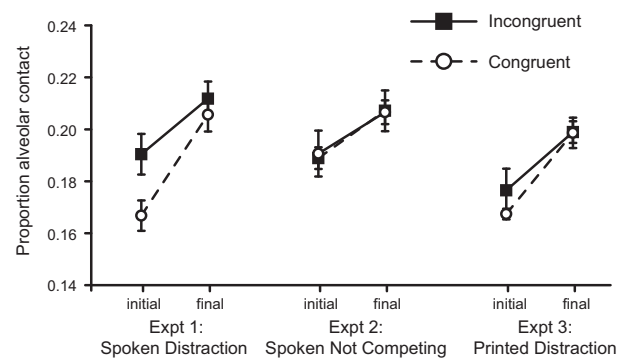


Table 1.

	/k/ targets		/s/ targets	
	<u>Initial</u>	<u>Final</u>	<u>Initial</u>	<u>Final</u>
Congruent	0.009	0.009	0.324	0.402
Incongruent	0.018	0.015	0.363	0.409

Table 2.

	<u>/k/ targets</u>		<u>/s/ targets</u>	
	<u>Initial</u>	<u>Final</u>	<u>Initial</u>	<u>Final</u>
Congruent	0.001	0.001	0.380	0.412
Incongruent	0.002	0.003	0.376	0.412

Table 3.

	/k/ targets		/s/ targets	
	<u>Initial</u>	<u>Final</u>	<u>Initial</u>	<u>Final</u>
Congruent	0.004	0.004	0.331	0.393
Incongruent	0.001	0.003	0.352	0.395