Original Paper

Phonetica 2012;69:109–123
DOI: 10.1159/000336117

Received: December 8, 2011
Accepted: January 21, 2012

Talking while Chewing: Speaker Response to Natural Perturbation of Speech

Connor Mayer Bryan Gick

University of British Columbia, Vancouver, B.C., Canada

Abstract

This study looks at how the conflicting goals of chewing and speech production are reconciled by examining the acoustic and articulatory output of talking while chewing. We consider chewing to be a type of perturbation with regard to speech production, but with some important differences. Ultrasound and acoustic measurements were made while participants chewed gum and produced various utterances containing the sounds /s/, /ʃ/, and /r/. Results show a great deal of individual variation in articulation and acoustics between speakers, but consistent productions and maintenance of relative acoustic distances within speakers. Although chewing interfered with speech production, and this interference manifested itself in a variety of ways across speakers, the objectives of speech production were indirectly achieved within the constraints and variability introduced by individual chewing strategies.

Introduction

Although the physiological processes of chewing and swallowing have been studied extensively, there have been few attempts to examine their interaction with speech, another primary function of the vocal tract. The present research aims to examine what happens when these two behaviours conflict by looking at the acoustic and articulatory effects of chewing on talking.

Talking while chewing can be viewed as a type of perturbation, though it differs in several important respects from perturbations previously studied. Past studies have used devices such as bite blocks [McFarland and Baum, 1995; McFarland et al., 1996], artificial palates [Honda et al., 2002], or loads on the lower jaw or lip [Abbs and Gracco, 1984; Kelso et al., 1984; Munhall et al., 1994]. These studies have collectively found articulatory, acoustic, and perceptual effects on speech production that underscore both the spatial-motor and acoustic-auditory goals of speech. Unlike the above methods, which introduce externally controlled perturbations outside the realm of speakers'

B.G. also at Haskins Laboratories, New Haven, Conn., USA.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2012 S. Karger AG, Basel
0031–8388/12/0693–0109
$38.00/0
Accessible online at:
www.karger.com/pho

Connor Mayer
Department of Linguistics, Totem Field Studios
2613 West Mall
Vancouver, BC V6T 1Z4 (Canada)
Tel. +1 778 230 3205, E-Mail connorm@interchange.ubc.ca

typical experience and are controlled externally to the speaker, the perturbation caused by talking while chewing is experienced frequently by speakers and is under speakers' control. In addition, the location, consistency and configuration of the food bolus changes over time, requiring constant readjustment, and it interferes with both the movement of the articulators and the shape of resonating cavities in the mouth. Speaking has been shown to share many kinematic features with chewing [Ostry and Flanagan, 1989]. The interplay between chewing and speaking puts conflicting demands on resources shared between these two tasks, such that a resolution of this conflict must be sensitive not only to the demands of speech production, but also to the objectives and strategies of chewing.

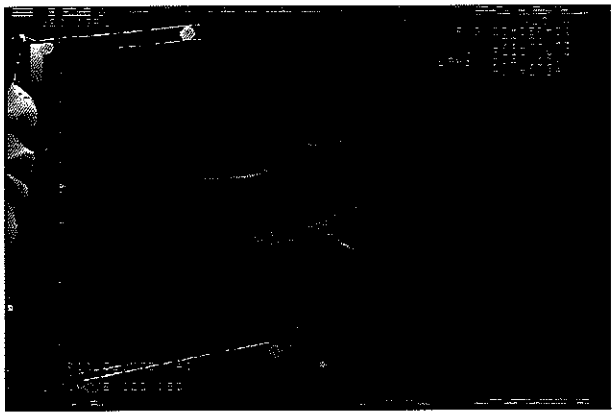
It has been shown that individuals vary greatly in their chewing patterns, with differences including chewing rate, duration, proportional muscle use, work rate, and the number of times the bolus is swapped from side to side, among other parameters [Brown et al., 1994b]. Two broad questions that arise from this variability among individual chewing strategies are whether the processes of speech production will constrain the variability of chewing across speakers, and whether chewing will introduce variability into the somewhat more regular processes of speech production. This paper focuses primarily on the latter question. The dynamic and constantly changing nature of chewing as a type of speech perturbation suggests that speakers may exhibit more variability in articulation and acoustics when speaking while chewing due to the changing position of the bolus. Further, the existence of individual chewing patterns suggests that at least part of the variability introduced may be contingent on a speaker's particular chewing behaviour. Thus we expect to see speakers exhibiting a range of individual variation, with articulatory and acoustic output reflecting the variability found in chewing behaviour. It is also possible that the individual variation in chewing behaviour is below the threshold that would seriously compromise typical speech production.

In this paper we aim primarily to address the question of how chewing perturbation influences the normal processes of speech production, looking at both articulation and acoustics. Previous research has shown that both articulatory and acoustic goals are important in speech production, at least for some sounds [Perkell et al., 2004; Ghosh et al., 2010]. These studies describe the sounds /s/ and /ʃ/ in terms of both articulatory and acoustic goals: it is likely that the perturbation produced by chewing will interfere with both articulatory and acoustic output of these sounds. Other work [Guenther et al., 1999; Perkell, in press] suggests that American English /r/ is likely to be similarly sensitive to perturbation caused by chewing.

In addition to the question of how the articulatory and acoustic characteristics of a given sound are impacted by chewing, a further aim of this study is to examine the effects of chewing on the acoustic relationship between sounds that are distinguished on the same acoustic dimension. Particularly we are interested in whether acoustic output is matched in an absolute sense (i.e. whether acoustic output when chewing falls within the range exhibited during normal speech) or in a relative sense (i.e. acoustic output does not fall within the range exhibited during normal speech but the relative distances between sounds in acoustic space are still preserved).

Methods

A study was conducted to observe the perturbation of speech by simultaneous chewing. Participants were seated in a modified dentist's chair with a cupped head support. An Aloka Prosound SSD-5000 ultrasound machine with a 180-degree EV probe stabilized by a mechanical arm was used



Color version available online

Fig. 1. Profile video superimposed over ultrasound video of P5 producing /r/ while chewing. Tongue tip is on the right.

to record midsagittal images of the tongue. Although midsagittal ultrasound only captures a small part of the potential differences in articulation introduced while chewing, it has the advantage of imaging at a relatively high frame rate along the whole length of the tongue, and is sufficient to provide an indication of many important distinctions in speech articulation. Differences off midline and in the size and shape of resonant cavities were not captured by this method. The probe was fitted with a 10-mm gel spacer to ensure participants' comfort while chewing. This made imaging of the tongue tip somewhat difficult, which may have resulted in increased variability in the tongue tip region. Profile video of the participants' faces in front of a blue screen was taken using a Sony Mini-DV Handicam. Participants wore a pair of sunglasses; two sticks covered in blue construction paper, each with two pink dots affixed, were attached to the sunglasses and to the probe. These dots showed the position of the head relative to the probe, allowing for correction of head movement once tongue shapes had been traced [Mielke et al., 2005]. The Chromakey feature of a Videonics MXPro DV video mixer was used to combine the two video channels, superimposing the participant's face and the dot positions over the ultrasound video, as shown in figure 1. Audio was recorded using a Sennheiser MK66 short shotgun microphone. The video and audio signals were fed through a Canopus ADVC-110 advanced digital video recorder into a MacPro computer, where the combined video and audio were captured using iMovie 8.0.6.

The stimuli consisted of the carrier phrase 'I'm a ___', followed by one of three words containing the phonemes of interest: 'saw', 'shaw', or 'raw'. The sounds /s/, /ʃ/, and /r/ were chosen because all three have fairly anterior oral non-closure constrictions (/r/ has multiple places of articulation; we refer here to the alveopalatal constriction) which are likely to be more susceptible to interference from the bolus, and all three have been described using relatively straightforward unidimensional acoustic measures that can be taken during the duration of the sound as a consequence of the incomplete articulatory closure: the frequency of spectral peaks for /s/ and /ʃ/ [Johnson, 2003], and a depressed F3 for /r/ [Delattre and Freeman, 1968]. This allows us to examine specifically the realization of these features while passing over other less contrastively important acoustic and articulatory consequences of the bolus. /s/ and /ʃ/ were also chosen because they are similar sounds that are largely distinguished on the same acoustic dimension, allowing for examination of the acoustic relationship between the two. Finally, as mentioned above, previous studies have shown that /s/ and /ʃ/ [Perkell et al., 2004; Ghosh et al., 2010] and /r/ [Guenther et al., 1999; Perkell, in press] rely on both acoustic and articulatory goals, and hence are likely to be sensitive to the acoustic and articulatory perturbation introduced by chewing.

There were two conditions: an experimental (With-Bolus) condition, in which speakers produced the stimuli while chewing a large bolus consisting of four pieces of Wrigley's^{TM} spearmint or fruit gum, and a control (No-Bolus) condition. Stimuli were presented in five blocks for each condition. Each block contained four repetitions of each word. The word orders within blocks were generated randomly using a computer program, but produced word orders lacking a relatively even distribution of sounds across the block were discarded. The first block for each condition was discarded, resulting in 32 tokens for each word (16 in each condition). Blocks alternated between the No-Bolus and With-Bolus conditions. Prior to the first With-Bolus condition, participants were given time to sufficiently chew the gum, and participants used the same mass of gum for the duration of the experiment, placing it in a dish during the control blocks. These steps ensured that the volume and consistency of the gum remained the same for each block. During the With-Bolus conditions, participants were instructed to continue chewing for the duration of the block, as though they were having a conversation over dinner. Speakers who paused their chewing to speak were reminded to chew continuously; 16 native speakers of Canadian English participated in the experiment.

All acoustic measurements were done using Praat 5.1.12. The durational boundaries and midpoints of /s/ and /ʃ/ were marked by hand, and acoustic centre of gravity (COG) measurements were made with a 30-ms window around the midpoints using Praat's COG measurement tool. COG is a measure of how high the frequencies in a spectrum are on average, and has proved to be useful for characterizing fricatives [Forrest et al., 1988]. For /r/, F3 was measured by hand at F3 minima. Using ELAN 3.7.2-1, still frames of the combined ultrasound and profile video were extracted from the midpoints for /s/ and /ʃ/ and from the F3 minima for /r/. The Palatoglossatron application [Baker, 2006] was used to trace tongue shapes and align the tracings using the Palatron algorithm to compensate for head movement [Mielke et al., 2005]. Output from Palatoglossatron was post-processed using the application Peterotron to compensate for a bug in the Palatoglossatron implementation of the Palatron algorithm.

All statistical tests were done using R 2.9.1, with a minimum significance level of 0.05. Repeated measure ANOVAs were done on pooled participant data to compare COG measurements for /s/ and /ʃ/ and F3 measurements for /r/ across conditions with participant as a random effect. Subsequently Welch's two-tailed t tests were used to examine individual differences across conditions. Several distributions in both conditions were found to be non-normal by visual inspection of the histogram and a Shapiro-Wilks normality test: in these cases, a Wilcoxon signed-rank test was used instead. This did not result in any differences in significance, however, so only the results of the t tests are reported.

An F test was used to examine differences in variance between conditions for each sound within individual participants. As mentioned above, certain distributions were found to be non-normal. Because the F test is quite sensitive to non-normality, in these cases one of two approaches was taken: in certain distributions (P3 /ʃ/ With-Bolus and /r/ No-Bolus, P8 /ʃ/ With-Bolus, P12 /s/ No-Bolus, P12 /ʃ/ With-Bolus, and P14 /s/ With-Bolus) single outliers were removed to make the distribution normal. Others (P4 /r/ With-Bolus, P7 /r/ No-Bolus, and P11 /r/ No-Bolus) could not be made normal by the removal of outliers, and so the results of the F tests are omitted. A Welch's two-tailed t test was used to see whether /s/ and /ʃ/ were significantly different in COG in both conditions.

Distributions of COG and F3 were examined in the bolus conditions to look for cases where more than one consistent response was exhibited by speakers. Multiple distinct responses to the perturbation could be reflected as non-normal bi- or multimodal distributions in the acoustic output, assuming that differing responses did not result in identical acoustic output.

Smoothing Spline ANOVA tests [Davidson, 2006] were performed on the ultrasound tongue tracings. The two curves represent a total of 16 tongue curves for each condition that were pooled across blocks and repetitions (excluding the first block). The dashed lines surrounding the solid lines in the images represent 95% confidence intervals; this corresponds to an alpha of 0.05 similar to the tests on acoustic data. Regions of significant difference are represented by areas where the confidence intervals for the two curves do not overlap. Following Davidson [2006], we define tongue blade, dorsum, and root by dividing the tongue contour into three equal parts: the leftmost third corresponds to the blade, the middle to the dorsum, and the rightmost to the root.

Results

Acoustic Results

Differences in Means between Conditions across Participants. A significant difference in COG across conditions was found for /s/ ($p < 0.001$) but not for /ʃ/

Table 1. Mean COG/F3 in Hertz across conditions with results of t test

Participant	/s/ NB (COG)	/s/ B (COG)	p value	/ʃ/ NB (COG)	/ʃ/ B (COG)	p value	/r/ NB (F3)	/r/ B (F3)	p value
P1	*8,863*	*8,138*	*<0.01*	*6,387*	*5,717*	*<0.05*	*1,972*	*1,868*	*<0.001*
P2	7,088	6,763	0.42	4,889	4,920	0.9	*1,588*	*1,532*	*<0.05*
P3	8,422	8,160	0.32	4,758	4,628	0.75	*1,307*	*1,408*	*<0.01*
P4	7,825	7,546	0.14	6,787	6,643	0.62	*1,437*	*1,335*	*<0.01*
P5	*8,242*	*7,643*	*<0.05*	*5,895*	*5,454*	*<0.01*	*1,798*	*1,638*	*<0.001*
P6	*6,984*	*7,731*	*<0.01*	*5,352*	*5,687*	*<0.05*	1,521	1,487	0.11
P7	*7,106*	*6,308*	*<0.05*	4,581	4,404	0.43	2,011	1,962	0.45
P8	9,047	8,465	0.1	5,172	5,598	0.11	1,888	1,913	0.38
P9	9,599	9,259	0.16	*7,130*	*8,000*	*<0.01*	*1,501*	*1,397*	*<0.05*
P10	*8,507*	*7,566*	*<0.0001*	*5,989*	*5,453*	*<0.01*	*1,443*	*1,366*	*<0.01*
P11	*7,736*	*8,323*	*<0.05*	*6,531*	*6,096*	*<0.01*	1,635	1,637	0.95
P12	*10,389*	*8,606*	*<0.0001*	6,221	6,259	0.92	*1,870*	*1,980*	*<0.01*
P13	*10,676*	*9,555*	*<0.01*	6,935	6,435	0.17	*1,830*	*1,757*	*<0.05*
P14	9,006	8,943	0.77	6,290	6,477	0.25	1,890	1,878	0.8
P15	8,883	8,351	0.15	5,776	5,630	0.62	*1,833*	*1,780*	*<0.05*
P16	7,050	7,156	0.73	5,899	5,781	0.71	*1,221*	*1,154*	*<0.01*

Italics indicate significant differences in mean. NB = No-Bolus condition; B = With-Bolus condition.

($p = 0.22$). COG was lower in the With-Bolus condition for both /s/ (8,464 Hz No-Bolus vs. 8,032 Hz With-Bolus) and /ʃ/ (5,912 Hz No-Bolus vs. 5,824 Hz With-Bolus). A marginally significant difference in F3 was found for /r/ across conditions ($p = 0.08$). F3 for /r/ was lower in the With-Bolus condition (1,631 Hz) than in the No-Bolus condition (1,672 Hz).

Differences in Means between Conditions within Participants. Significant differences in COG for /s/ were found in half the participants: participants P1, P5, P6, P7, P10, P11, P12, and P13. Of these 8 participants, all except P6 and P11 showed a lower COG for /s/ in the With-Bolus condition. For /ʃ/, significant differences in COG were found in 6 participants: 1, 5, 6, 9, 10, and 11, with all except P6 and P9 displaying lower COG in the With-Bolus condition. With the exception of P9, all participants who displayed differences in /ʃ/ also displayed differences in /s/. Significant differences for F3 minima in /r/ were found in P1, P2, P3, P4, P5, P9, P10, P12, P13, P15, and P16, with all 11 participants except for P3 and P12 displaying significantly lower F3 minima in the With-Bolus conditions. /r/ was the most common sound to display acoustic interference from the bolus, followed by /s/, then /ʃ/. When the presence of the bolus did result in significantly different acoustic output, it tended to be lower COG in /s/ and /ʃ/, and lower F3 minima in /r/. These results are displayed in table 1.

Differences in Variance across Conditions within Participants. P1, P2, P3, P4, P5, P8, P9, and P16 showed significant differences in variance of COG for /s/ between conditions. P1, P2, P12, P13, and P16 showed significant differences in variance of COG for /ʃ/ between conditions. P1, P5, and P9 showed significant differences in variance of F3 minima for /r/ between conditions. These results are shown in table 2.

Differences between /s/ and /ʃ/ in Individual Participants within Conditions. All participants displayed significant differences between /s/ and /ʃ/ in the No-Bolus

Table 2. Standard deviations in Hertz for With-Bolus (B) and No-Bolus (NB) conditions with results from F tests

Participant	/s/ B (COG)	/s/ NB (COG)	p value	/ʃ/ B (COG)	/ʃ/ NB (COG)	p value	/r/ B (F3)	/r/ NB (F3)	p value
P1	*807*	*468*	*<0.05*	*1,129*	*406*	*<0.001*	*44*	*82*	*<0.05*
P2	*1,519*	*379*	*<0.0001*	*902*	*418*	*<0.01*	64	52	0.44
P3	*899*	*526*	*<0.05*	568	499	0.63	104	162	0.24
P4	*662*	*291*	*<0.01*	858	742	0.58	49	58	N/A
P5	*897*	*523*	*<0.05*	514	317	0.07	*147*	*74*	*<0.05*
P6	604	665	0.72	460	456	0.97	71	44	0.08
P7	918	841	0.74	728	496	0.15	198	139	N/A
P8	*655*	*1,152*	*<0.05*	506	532	0.85	73	67	0.76
P9	*880*	*269*	*<0.0001*	716	863	0.48	*140*	*46*	*<0.0001*
P10	521	452	0.59	509	424	0.49	89	55	0.07
P11	654	566	0.58	441	309	0.18	85	89	N/A
P12	946	633	0.23	*887*	*353*	*<0.01*	103	102	0.99
P13	971	764	0.36	*1,271*	*632*	*<0.05*	98	96	0.93
P14	784	371	0.06	468	429	0.74	150	110	0.25
P15	931	1,081	0.58	830	810	0.93	65	69	0.83
P16	*605*	*1,043*	*<0.05*	*1,120*	*563*	*<0.05*	72	47	0.11

Italics indicate significant differences in variance between conditions.

condition, and all participants, even those who displayed significant differences in COG in one or both of /s/ and /ʃ/ across conditions, showed significant differences between /s/ and /ʃ/ in the With-Bolus condition. The differences were highly significant, with all p values in the No-Bolus condition being less than 0.001 and all p values for differences in the With-Bolus condition being less than 0.001 except for participant 4 ($p < 0.01$).

Distributions. In the With-Bolus condition, only P3 /ʃ/, P4 /r/, P8 /ʃ/, P12 /ʃ/, and P14 /s/ displayed non-normal distributions: of these, all except P4 /r/ were made non-normal by the presence of one extreme outlier. Only P4 displays a truly bimodal distribution: however, the F3 of the individual tokens varies between the areas of the two peaks seemingly at random throughout the course of the experiment. While this does not in itself preclude the existence of two responses to a frequently changing bolus position, the small number of tokens makes it possible that this bimodal distribution occurred at random.

Articulatory Measurements

Significant differences in tongue position between the With-Bolus and No-Bolus conditions for /s/ were found in 5 participants: P1 in the tongue root, P7 in the tongue root, P8 in the tongue blade and dorsum, P9 in the tongue blade and root, and P16 in the tongue root. For /ʃ/, significant differences were found in 8 participants: P1 in the tongue root, P2 in the tongue blade and root, P4 in the tongue blade, P5 in the tongue blade, P6 in the tongue root, P9 in the tongue blade and root, P13 in the tongue blade, and P16 in the tongue root. Significant differences in tongue position for /r/ were found in 4 participants: P1 in the tongue root, P12 in the tongue tip, P13 in the tongue tip, and P16 in the tongue blade.

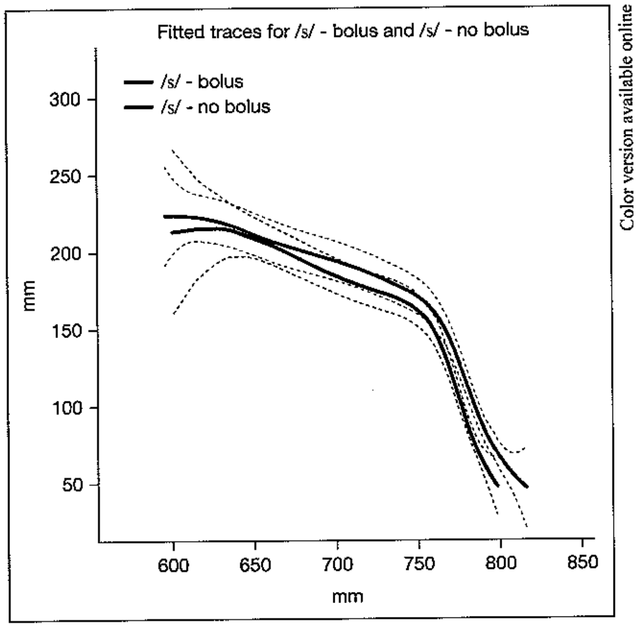


Fig. 2. Type A articulation of /s/ by P1. Tongue tip is on the left.

Following the treatment of /r/ in Delattre and Freeman [1968], it is possible to create several broad categories of variation in the With-Bolus conditions for /s/, /ʃ/, and /r/. These categories focus only on statistically significant areas of difference in tongue shape and are described in the next three sections. The images shown are averaged tongue tracings. The tongue tip is on the left for all images.

Variants for /s/. Articulatory differences in the With-Bolus condition for /s/ can be classified into three rough types. Type A, an example of which is shown in figure 2, is characterized by a significantly advanced tongue root, and is displayed by P1, P7, and P8. These participants also displayed a tendency towards a lower blade and dorsum, but this was only significant for P8. Type B, characterized by a raised and more posterior tongue blade, is displayed by P9 in figure 3. Type C, a retracted tongue root, is displayed by P16 in figure 4.

Variants for /ʃ/. For /ʃ/, participants' articulations in the With-Bolus condition can be divided into four types. Type A, displayed by P2, P4, P9, and P13, is characterized by a higher and more anterior tongue blade. An example is shown in figure 5. Type B is characterized by a retracted tongue root. This was displayed by P6 and P16, and an example is shown in figure 6. Type C, exhibited by P5 in figure 7, is characterized by a lowered tongue blade. Finally, type D is characterized by an advanced tongue root, and is shown by P1 in figure 8.

Variants for /r/. The articulations for /r/ in the With-Bolus condition can be broken down into two categories. The first, category A, is characterized by an advanced tongue root. Only P1 displayed this pattern, as shown in figure 9. Category B, displayed by P12, P13, and P15, is characterized by a lowered tongue blade for both bunched and retroflex productions. An example is shown in figure 10.

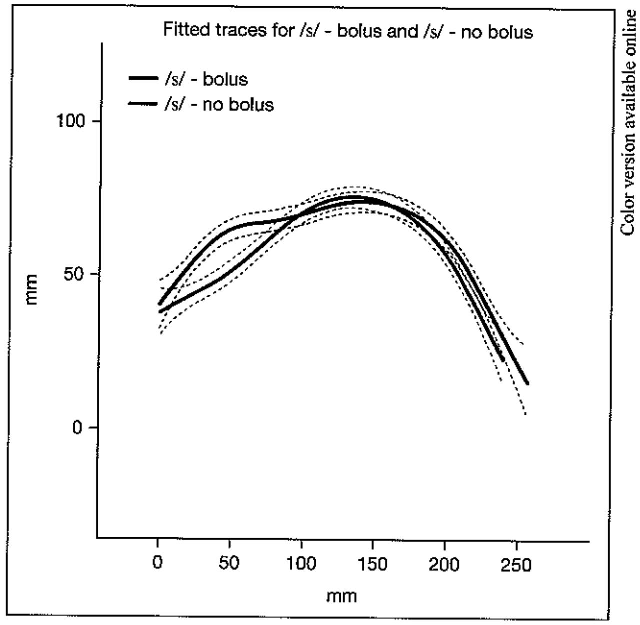


Fig. 3. Type B articulation of /s/ by P1. Tongue tip is on the left.

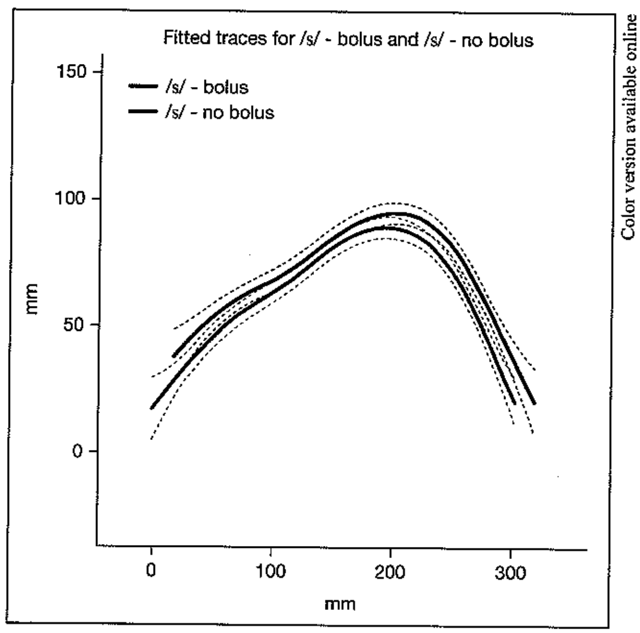


Fig. 4. Type C /s/ articulation by P16. Tongue tip is on the left.

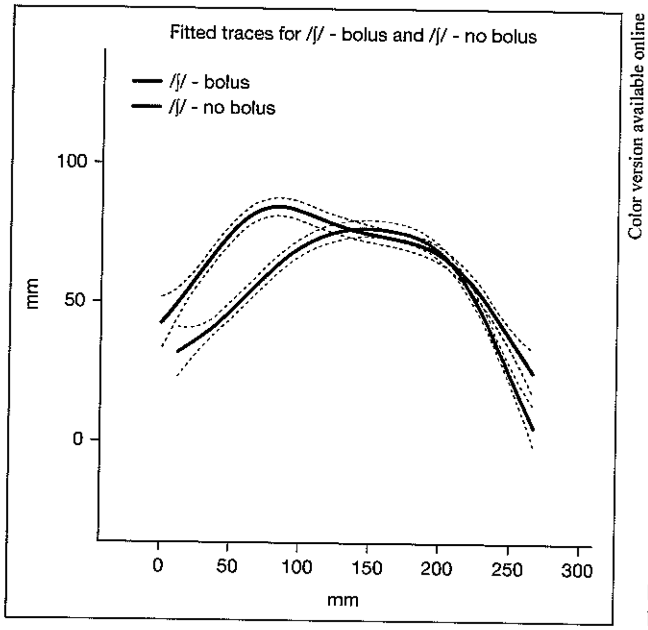


Fig. 5. Type A /ʃ/ articulation by P9. Tongue tip is on the left.

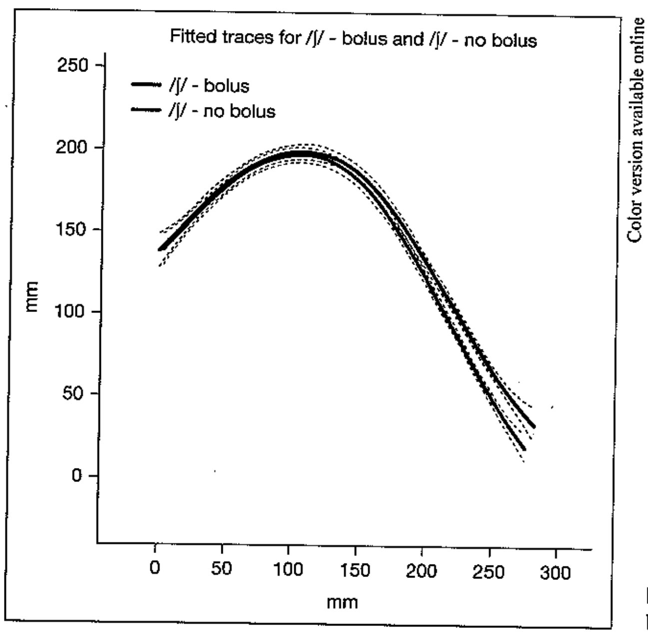
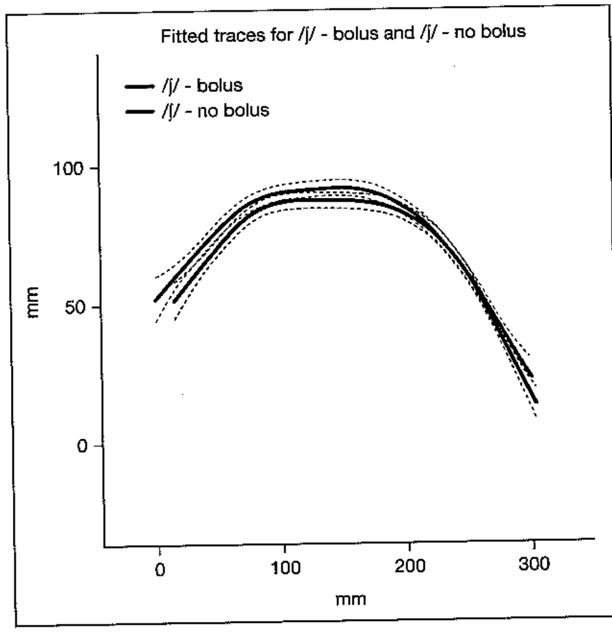
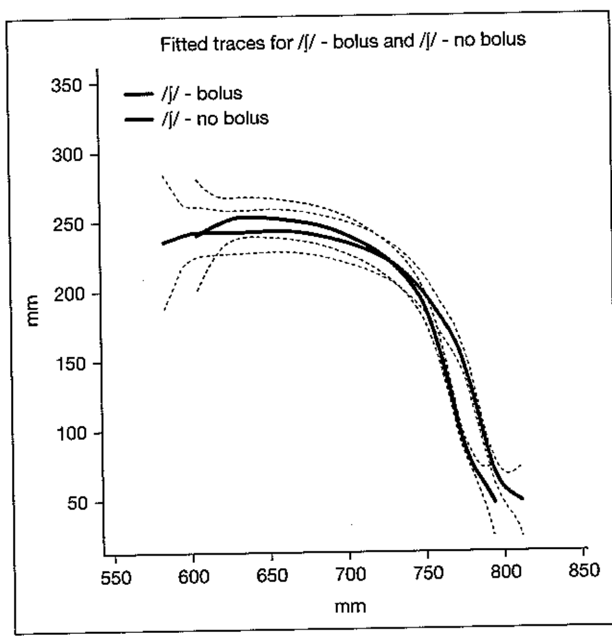


Fig. 6. Type B /ʃ/ articulation by P6. Tongue tip is on the left.



Color version available online

Fig. 7. Type C /ʃ/ articulation by P5. Tongue tip is on the left.



Color version available online

Fig. 8. Type D /ʃ/ articulation by P1. Tongue tip is on the left.

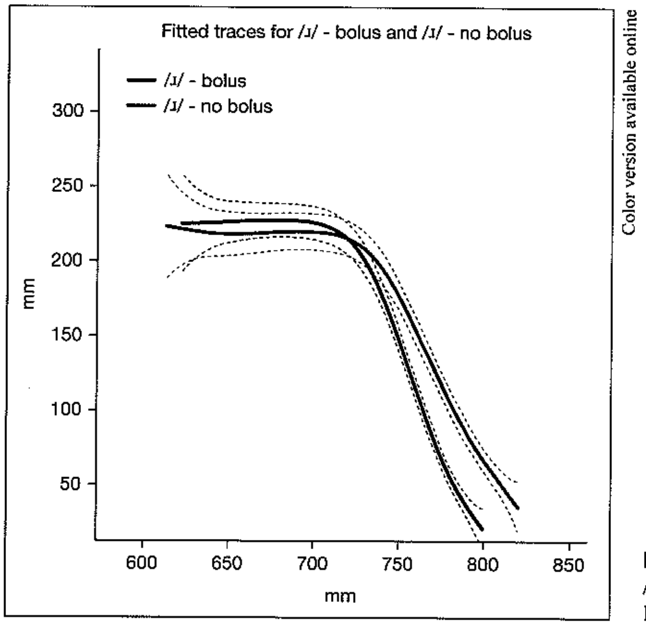


Fig. 9. Type A articulation of /r/ by P1. Tongue tip is on the left.

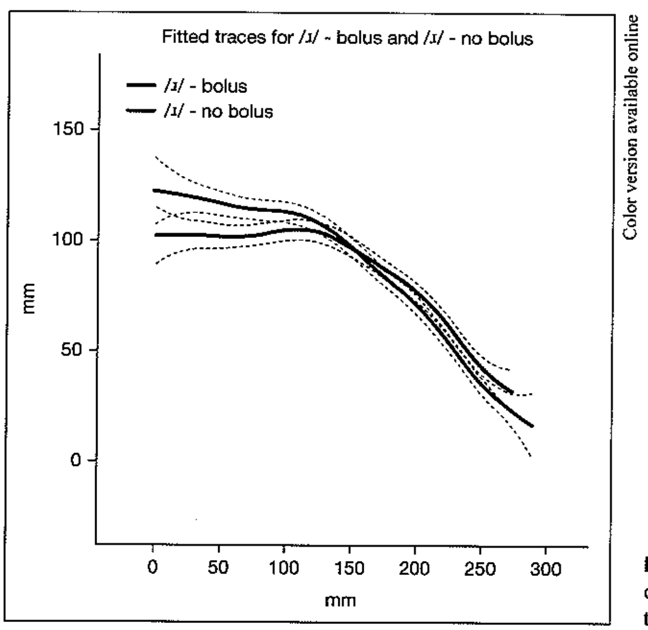


Fig. 10. Type B articulation of /r/ by P13. Tongue tip is on the left.

Discussion

Maintenance of Acoustic Output and Distinctions across Conditions

Pooled participant data showed a significant difference in acoustic output for /s/ and a marginally significant difference for /r/, while no significant difference was seen for /ʃ/. This indicates that /s/ and to a lesser degree /r/ are more susceptible to interference from the bolus than /ʃ/.

Turning to individual differences, over half the participants showed acoustic output in the With-Bolus condition that fell within the range exhibited in the No-Bolus condition for /s/ and /ʃ/. Typically participants who displayed a significant difference in COG for /ʃ/ also showed a difference in /s/. The only exception to this is P9. This suggests that if the bolus interferes with production of a more posterior sound such as /ʃ/ it is also likely to have affected more anterior constrictions. Participants were less successful at maintaining absolute acoustic output for /r/, with only 5 of the 16 showing no differences between conditions.

The majority of participants produced lower COG and F3 in the With-Bolus condition. Because COG [Johnson, 2003] and F3 [Espy-Wilson et al., 2000] are negatively correlated with the cavity size anterior to the constriction, this suggests that the gum in these cases may have resulted in a larger anterior cavity. Eight participants showed significant differences in two or more of the sounds. Of these 8, 5 were consistent in the direction of change across sounds, while 3 were not, suggesting that a larger anterior cavity was not always present across sounds. We interpret this as being indicative of varying chewing strategies both within and across individuals.

Despite differences in acoustic output between conditions for /s/ and /ʃ/ in some participants, there were no cases where the distinction between /s/ and /ʃ/ was lost in the With-Bolus condition. This suggests that even when the presence of the bolus forced changes in acoustic production, efforts were still made to maintain relative distinctions in acoustic space between /s/ and /ʃ/. Relative acoustic distinctions were maintained even as absolute goals were not. For /r/, despite the general lack of success in preserving typical acoustic production, the differences typically consisted of the most salient feature, a low F3, being maintained and even emphasized. Thus, although individual chewing strategies introduced a significant amount of variability into speech production, the goals of chewing did not essentially compromise the goals of speech production.

Acoustic Variability across Conditions

Differences in variability across conditions, when present, tended to be found in /s/ and /ʃ/. It is possible that these anterior constrictions were more susceptible to interference from the gum, resulting in increased variance. This is related to the finding that /r/ displayed the fewest instances of articulatory and variance differences, but the most differences in F3, suggesting that the presence of a bolus interferes with resonating cavities, but does not actually affect normal articulation of /r/ in most cases. A higher standard deviation in the bolus condition for participants with significant differences in variance was displayed by 6 of the 8 participants for /s/, all participants for /ʃ/, and 2 of the 3 participants for /r/. This indicates that the presence of a constantly changing bolus can increase the variability of acoustic output, but the effect does not seem to be consistent across participants: the fact that 3 participants displayed significantly lower variance in the bolus condition may be evidence that the chewing strategies for these participants resulted in greater constraints upon the articulatory processes of speech production. There is no clear relationship between those participants who displayed significant differences in COG or F3 across conditions and those that displayed significant differences in variance.

Individual Strategies

The normality of the distributions of COG or F3 minima in the With-Bolus condition suggests that participants attempt to maintain stable acoustic output despite the presence of gum, and are for the most part successful, although the presence of outliers in several participants shows that the changing position of the bolus can occasionally have unexpected acoustic consequences. Whether or not this normality is the result of a single consistent compensatory strategy or multiple similarly effective strategies is difficult to say with the current data. Based on the classification of chewing behaviour into several distinct types by Brown et al. [1994a], it is likely that speakers' chewing strategies interact in a consistent manner with their speech production. In the majority of cases in this study, absolute acoustic goals were not met but the With-Bolus condition exhibited a normal distribution. This seems to suggest that when a conflict occurs, a stable chewing strategy will trump stable acoustic output. Only P4 deviated from this normality, showing a bimodal distribution for the F3 minima in /r/. As mentioned above, this is more likely to be a result of the small number of tokens rather than true bimodality.

Maintenance of Articulatory across Conditions

Participants tended to show more consistency across conditions for midsagittal articulation than for acoustics, with only 17 total articulatory differences to 25 differences in mean COG or F3 across conditions. Furthermore, participants did not seem to display robust trading relations between articulation and acoustics [Guenther et al., 1999]: of a possible 48 cases, only 7 were found where there was an articulatory difference but no significant difference in the acoustic measure (and of course in these cases there could have been other acoustic differences not measured in the present study). The prevalence of acoustic differences compared to articulatory ones could be due to the gum interfering with acoustics by altering the size of resonant cavities – including the buccal cavities, or by articulatory interference or compensation in areas of the vocal tract not imaged using the present ultrasound techniques.

Although it is possible to divide participants who displayed significant articulatory differences into various categories based on the location and type of the difference, there does not seem to be any clear correspondence between the type of articulatory difference observed and the differences seen in our acoustic measures.

Conclusions

This study indicates that /s/ and /r/ are more susceptible to acoustic interference from chewing while speaking than /ʃ/. Despite this interference, however, participants were fairly skilled at maintaining typical acoustic output, if not in an absolute sense then in a relative sense that preserved acoustic differences. Typical articulation on the midsagittal plane was also generally maintained. Perhaps the most striking result of this study, however, is the high degree of variation between individuals. No two participants displayed identical patterns: there does not seem to be any relationship between the differences across conditions in mean COG or F3, variance, or midsagittal articulation, in that the presence of one could not predict the presence of another in a consistent way across participants. While there were certain broad tendencies, such as the With-Bolus condition showing lower COG and F3 minima or certain articulatory differences exhibited by several participants, there were frequent exceptions. This suggests that chewing will introduce variability in speech production when the two behaviours conflict, and that the exact effect on acoustic and articulatory output varies with individual chewing strategies: certain participants like P1 showed significantly impaired acoustics and articulation for all segments, while others, such as P14, showed few or even no differences. Despite this high degree of cross-participant variability, productions from individual speakers were quite stable, as evidenced by the unimodal distribution of the acoustic output in the With-Bolus condition. Thus it seems that the cross-speaker differences are not simply the result of the variable nature of chewing but rather that of varying chewing strategies that are consistent within individuals. Speakers still maintain relative distinctions between sounds even when chewing forces the abandonment of absolute distinctions. This suggests that speech production is not completely compromised in this situation: rather its goals are indirectly achieved within the constraints introduced by an individual's chewing strategy.

There remain many questions that should be addressed by future research. Foremost of these is a more thorough examination of the relationship between individuals' specific chewing patterns and their compensatory responses to perturbation caused by chewing. Broader articulatory measures such as coronal ultrasound, X-ray or fast MRI could also give a more complete picture of the articulatory perturbation and compensation that occurs while chewing, as could biomechanical modeling. In addition, data on the exact nature of the variability found in chewing would be quite valuable in regard to this question, and should be pursued in future research.

References

Abbs, J.H.; Gracco, V.L.: Control of complex motor gestures: Orofacial muscle responses to load perturbations of the lip during speech. J. Neurophysiol. *51:* 705–723 (1984).

Baker, A.: Palatoglossatron 1.0. http://dingo.sbs.arizona.edu/~apilab/pdfs/pgman.pdf (2006).

Brown, W.E.; Langley, K.R.; Martin, A.; Macfie, H.J.H.: Characterisation of patterns of chewing behaviour in human participants and their influences on texture perception. J. Texture Stud. *25:* 455–468 (1994a).

Brown, W.E.; Shearn, M.; Macfie, H.J.H.: Method to investigate differences in chewing behaviour in humans. II: Use of electromyography during chewing to assess chewing behaviour. J. Texture Stud. *25:* 17–31 (1994b).

Davidson, L.: Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. J. acoust. Soc. Am. *120:* 407–415 (2006).

Delattre, P.; Freeman, D.: A dialect study of American r's by X-ray motion picture. Linguistics *44:* 29–68 (1968).

Espy-Wilson, C.Y.; Boyce, S.E.; Jackson, M.; Narayanan, S.; Alwan, A.: Acoustic modeling of American English /r/. J. acoust. Soc. Am. *108:* 343–356 (2000).

Forrest, K.; Weismer, G.; Milenkovic, P.; Dougall, R.N.: Statistical analysis of word-initial voiceless obstruents: Preliminary data. J. acoust. Soc. Am. *84:* 115–123 (1988).

Ghosh, S.S.; Matthies, M.L.; Maas, E.; Hanson, A.; Tiede, M.; Ménard, L.; Guenther, F.H.; Lane, H.; Perkell, J.S.: An investigation of the relationship between sibilant production and somatosensory and auditory acuity. J. acoust. Soc. Am. *128:* 3079–3087 (2010).

Guenther, F. H.; Espy-Wilson, C.Y.; Boyce, S.E.; Matthies, M.L.; Zandipour, M.; Perkell, J.S.: Articulatory tradeoffs reduce acoustic variability during American English /r/ production. J. acoust. Soc. Am. *105:* 2854–2865 (1999).

Honda, M.; Fujino, A.; Kaburagi, T.: Compensatory responses of articulators to unexpected perturbation of the palate shape. J. Phonet. *30:* 281–302 (2002).

Johnson, K.: Acoustic and auditory phonetics; 3rd ed. (Blackwell, Oxford 2003).

Kelso, J.A.S.; Tuller, B.; Vatikiotis-Bateson, E.; Fowler, C.A.: Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinate structures. J. exp. Psychol. hum. Perception Performance, *10:* 812–832 (1984).
McFarland, D.; Baum, S.: Incomplete compensation to articulatory perturbation. J. acoust. Soc. Am. *97:* 1865–1873 (1995).
McFarland, D.; Baum, S.; Chabot, C.: Speech compensation to structural modifications of the oral cavity. J. acoust. Soc. Am. *100:* 1093–1104 (1996).
Mielke, J.; Baker, A.; Archangeli, D.; Racy, S.: Palatron: a technique for aligning ultrasound images of the tongue and palate. Coyote Papers *14:* 147–161 (2005). http://coyotepapers.sbs.arizona.edu/CPXIV/CPXIV%20p96–107%20Mielke-Baker-Archangeli-Racy.pdf.
Munhall, K.G.; Löfqvist, A.; Kelso, J.A.S.: Lip-larynx coordination in speech: Effects of mechanical perturbations to the lower lip. J. acoust. Soc. Am. *95:* 3605–3616 (1994).
Ostry, D.J.; Flanagan, J.R.: Human jaw movement in mastication and speech. Archs oral Biol. *34:* 685–693 (1989).
Perkell, J.S.: Movement goals and feedback and feedforward control mechanisms in speech production. J. Neuroling. (in press). http://dx.doi.org/10.1016/j.jneuroling.2010.02.011.
Perkell, J.S.; Matthies, M.L.; Tiede, M.; Lane, H.; Zandipour, M.; Marrone, N.; Stockmann, E.: The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. J. Speech Lang. Hear. Res. *47:* 1259–1269 (2004).