ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Quantifying Articulatory Working Space in Individuals Surgically Treated for Oral Cancer With Electromagnetic Articulography

Thomas B. Tienkamp,[a,b,c] Teja Rebernik,[a,c,d] Bence M. Halpern,[d,e,f] Rob J. J. H. van Son,[d,e] Martijn Wieling,[a,c,g] Max J. H. Witjes,[b] Sebastiaan A. H. J. de Visscher,[b] and Defne Abur[a,c]

[a] Center for Language and Cognition Groningen, University of Groningen, the Netherlands [b] Department of Oral and Maxillofacial Surgery, University Medical Center Groningen, the Netherlands [c] Research School of Behavioral and Cognitive Neurosciences, University of Groningen, the Netherlands [d] Netherlands Cancer Institute, Amsterdam, the Netherlands [e] Amsterdam Center for Language and Communication, University of Amsterdam, the Netherlands [f] Multimedia Computing Group, Delft University of Technology, the Netherlands [g] Haskins Laboratories, New Haven, CT

ARTICLE INFO

Article History:
Received February 14, 2023
Revision received October 16, 2023
Accepted October 19, 2023

Editor-in-Chief: Peggy B. Nelson
Editor: Jun Wang

https://doi.org/10.1044/2023_JSLHR-23-00111

ABSTRACT

Purpose: The purpose of this study was to quantify sentence-level articulatory kinematics in individuals treated for oral squamous cell carcinoma (ITOC) compared to control speakers while also assessing the effect of treatment site (jaw vs. tongue). Furthermore, this study aimed to assess the relation between articulatory–kinematic measures and self-reported speech problems.
Method: Articulatory–kinematic data from the tongue tip, tongue back, and jaw were collected using electromagnetic articulography in nine Dutch ITOC and eight control speakers. To quantify articulatory kinematics, the two-dimensional articulatory working space (AWS; in mm^2), one-dimensional anteroposterior range of motion (AP-ROM; in mm), and superior–inferior range of motion (SI-ROM in mm) were calculated and examined. Self-reported speech problems were assessed with the Speech Handicap Index (SHI).
Results: Compared to a sex-matched control group, ITOC showed significantly smaller AWS, AP-ROM, and SI-ROM for both the tongue tip and tongue back sensor, but no significant differences were observed for the jaw sensor. This pattern was found for both individuals treated for tongue and jaw tumors. Moderate nonsignificant correlations were found between the SHI and the AWS of the tongue back and jaw sensors.
Conclusions: Despite large individual variation, ITOC showed reduced one- and two-dimensional tongue, but not jaw, movements compared to control speakers and treatment for tongue and jaw tumors resulted in smaller tongue movements. A larger sample size is needed to establish a more generalizable connection between the AWS and the SHI. Further research should explore how these kinematic changes in ITOC are related to acoustic and perceptual measures of speech.

Surgical treatment for oral squamous cell carcinoma (OSCC) results in lasting physiological changes that may complicate speech and/or swallowing (Kreeft et al., 2009). The tongue can become less mobile due to scar tissue and potential postoperative radiation therapy after treatment for tongue cancer (Jacobi et al., 2013; Kappert et al., 2019; Lazarus et al., 2014). Treatment for OSCC located on the jaw may also result in limited tongue mobility as malignancies occur rarely exclusively in the mandibular bone (Bak et al., 2010; de Groot et al., 2020).

Even though patients consistently rank speech as one of the top priorities during and following OSCC treatment, the articulatory–kinematic consequences of treatment have not yet been characterized in great detail

Correspondence to Thomas B. Tienkamp: t.b.tienkamp@rug.nl. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

(Tschiesner et al., 2013). Kinematic consequences have been mostly limited to studies investigating to what extent individuals treated for OSCC (ITOC) showcase similar or distinct movement patterns during speech compared to control speakers (Bressmann et al., 2007; Hagedorn et al., 2021; Stone et al., 2014). These studies have shown that tongue motion patterns differ between groups both qualitatively and quantitatively. On the one hand, both studies of Stone et al. (2014; $n = 3$) and Bressmann et al. (2007; $n = 12$) showed that surgical treatment for lateral tongue tumors resulted in more asymmetrical movement compared to typical speakers. On the other hand, Hagedorn et al. (2021; $n = 6$) showed that ITOC have less complex vocal tract shaping compared to typical speakers as ITOC had difficulty with differentially controlling distinct parts of the tongue. This further resulted in less movement in affected areas (e.g., less amplitude in the velar region if the tumor affected the base of the tongue). However, these studies did not directly quantify the absolute size of the movement (e.g., in mm) during speech.

The range of motion (ROM) of the tongue could be a promising measure to predict speech outcomes in ITOC, since a more mobile tongue (i.e., larger ROM values) usually leads to better and more intelligible speech posttreatment (Bressmann et al., 2004; Chepeha et al., 2016; Lam & Samman, 2013; Pauloski et al., 1998; van Dijk et al., 2016). However, studies that have assessed the ROM of the tongue following OSCC treatment have done so with nonspeech tasks, such as maximal tongue protrusion or lateral movement, quantified by Likert scales, ruler-based measurements, or three-dimensional camera tracking (Chepeha et al., 2016; Kappert et al., 2019; Lazarus et al., 2014; Speksnijder et al., 2011). Methods exploring maximum movement portray the maximum efforts in terms of anatomical capability, which may require different motor demands compared to speech tasks (Bressmann et al., 2004; Schuster & Stelzle, 2012). Existing work using articulatory–acoustic measures of speech suggests that the ROM of the tongue during speech is reduced as a consequence of OSCC treatment. Specifically, the second formant (F_2) of /i/ and center of gravity of /s/, both acoustic correlates of anteroposterior tongue movement, were shown to be reduced in ITOC (Acher et al., 2014; de Bruijn et al., 2009; Laaksonen et al., 2011; Takatsu et al., 2017; Tienkamp et al., 2023; Zhou et al., 2013). However, acoustic measures only provide indirect evidence of the underlying articulatory gestures. For example, ITOC may use the unaffected articulators (e.g., the jaw and/or lips) in a compensatory way to produce the same acoustic output (Hagedorn et al., 2022). This might complicate the acoustic–kinematic relationship that has been established for both typical speakers and individuals with dysarthria (Lee et al., 2017; Mefferd & Green, 2010). Consequently, articulatory–kinematic measures might become especially important in order to assess the ROM of the tongue in ITOC directly.

As the tongue is hidden in the oral cavity, measuring its ROM requires several methodological considerations. A survey among 292 speech-language pathologists (SLPs) in the United States showed that 88% estimated the ROM of the tongue based on visual clinical judgment and only 9% used a ruler to quantify the ROM of the tongue by measuring the distance from the upper lip to the tongue tip (TT; Husaini et al., 2014; Lazarus et al., 2014). The remaining 3% did not specify the measurement method. Despite the fact that the ROM is not usually quantified on a continuous scale in clinical practice, all survey respondents agreed that the ROM of the tongue is one of the most important predictors of speech and swallowing outcomes following OSCC treatment (Husaini et al., 2014). However, as van Dijk et al. (2016) note, physical contact of the ruler with the tongue is almost inevitable, which could impact its movement. A different, but subjective, method that is often used in research is to assess the ROM of the tongue on a 3-point Likert scale (e.g., Bressmann et al., 2004; Konstantinović & Dimić, 1998; Speksnijder et al., 2011). Granting that this eliminates the need for physical contact, specificity may be lost if only three categories are used and no measurements on a continuous scale are made. In response, a recently developed method used three-dimensional camera tracking in order to quantify the ROM of the tongue noninvasively on a continuous scale by attaching a tongue marker, which is subsequently tracked by three cameras (Kappert et al., 2019; van Dijk et al., 2016). One remaining disadvantage is that this setup can only capture tongue movements when the mouth is open as the cameras cannot track the marker if the tongue is not visible.

A method that is able to directly quantify tongue movements with both an open and closed mouth is electromagnetic articulography (EMA). EMA is a three-dimensional point-tracking system where sensors are attached to various parts of the articulators (e.g., the jaw, lips, and tongue) and tracked through a magnetic field. The advantage of using EMA is that it is able to track articulatory–kinematic motion with a very high spatial and temporal resolution, thus allowing for a precise assessment of the ROM of the tongue and other articulators during speech (Rebernik, Jacobi, Tiede, & Wieling, 2021; Savariaux et al., 2017). While EMA does not capture the movement of the entire tongue contour as each sensor can only capture one point on the tongue, it is able to measure the kinematic motion of distinct parts of the tongue (e.g., tongue front and tongue back).

The extent to which the tongue or other articulators move during speech is known as the articulatory working space (AWS). The AWS has been determined by calculating the global movement size during an entire speech utterance either in two-dimensional (mm^2) or in three-dimensional (mm^3) space using EMA. Movements of the articulators in each separate dimension, for example, in the anteroposterior or vertical dimension, are typically labeled as the anteroposterior ROM (AP-ROM) and superior–inferior ROM (SI-ROM). Several studies have investigated the AWS in individuals with and without speech impairments. By using EMA, Lee and Bell (2018) computed the two-dimensional AWS for the TT, tongue body, lower lip, and jaw in persons with amyotrophic lateral sclerosis (PwALS; n = 22) and typical speakers (n = 22) as the speakers read The Rainbow Passage (Fairbanks, 1960). Their results indicated that PwALS had a significantly lower AWS (in mm^2) for the TT and tongue body compared to typical speakers, whereas the AWS of the lower lip was significantly larger. PwALS also had significantly lower AP-ROM for the TT, whereas the AP-ROM of the lower lip and jaw was significantly larger, signaling some form of compensatory behavior. Moreover, the AWS size was related to both speech intelligibility and speech rate, such that bigger movements for both tongue sensors resulted in more intelligible and faster speech (Lee & Bell, 2018). Similarly, the AWS has been used to quantify the movement of the tongue, lips, and jaw in persons with Parkinson's disease (PwPD; n = 22), PwALS (n = 10), and controls (n = 20) using EMA (Weismer et al., 2012). Both dysarthric speaker groups were found to have smaller movement sizes in mm^2 for the tongue and lower lip compared to controls during a reading of the Hunter Passage (Crystal & House, 1982). Lastly, Kearney et al. (2017) used EMA to quantify the movement of the tongue and jaw in PwPD (n = 21) and controls (n = 20). Compared to controls, PwPD demonstrated smaller jaw movements in mm^3, but no differences in tongue movements in mm^3. Similar to the study of Lee and Bell, larger tongue movements were found to be associated with higher speech intelligibility. Thus, prior studies indicate that the AWS is able to capture kinematic differences between typical and dysarthric speech and seems to be associated with linguistically relevant aspects such as intelligibility and speech rate. Moreover, the AWS has relatively high ecological validity as it can be computed over running speech, making it a relevant tool to assess the consequences of OSCC treatment on tongue and jaw movement during speech.

Consequently, the purpose of the present study was to assess the articulatory function of ITOC compared to controls as quantified by both one-dimensional (AP-ROM and SI-ROM) and two-dimensional (the AWS) movement data for both the tongue and the jaw. Specifically, kinematic data recorded with EMA from nine ITOC and eight controls were used to examine the AWS, AP-ROM, and SI-ROM of the tongue and jaw.

The first aim of our study was to quantify the degree to which the AWS and the ROM in the anteroposterior and superior–inferior dimensions were reduced in ITOC compared to controls. We hypothesized that the AWS of ITOC would be smaller compared to controls based on previous work measuring nonspeech ROM (Bressmann et al., 2004; Chepeha et al., 2016; de Groot et al., 2020; Kappert et al., 2019; Speksnijder et al., 2011). Our second hypothesis was that ITOC would show a reduced AP-ROM as acoustic studies found reduced F_2 for /i/ and center of gravity for /s/ (i.e., acoustic correlates of reduced anteroposterior tongue movement; Acher et al., 2014; de Bruijn et al., 2009; Laaksonen et al., 2011; Takatsu et al., 2017; Tienkamp et al., 2023; van Dijk et al., 2016; Zhou et al., 2013). Based on this line of acoustic work, our final hypothesis was that we would not find reductions in SI-ROM.

The second aim of our study was to assess whether the two-dimensional movement size of the tongue and jaw in ITOC was related to the type of treatment they received. While our first aim was to assess group-level differences, our second aim assesses potential differences in a more fine-grained manner by looking at the effect of treatment site. Insight into which surgical procedures affect the AWS most severely may be helpful in shared decision making with regard to treatment options (Kappert et al., 2019). We predicted that the location of the tumor would affect the impacted articulator's mobility. That is, if an individual was treated for a tongue tumor or a tumor fixated on the jaw in proximity to the tongue, a smaller AWS of the tongue would be expected compared to control speakers. Similarly, we predicted a smaller AWS of the jaw for individuals treated for jaw tumors, but not for tongue tumors when compared to control speakers.

The third and final aim of the study was to estimate the degree to which two-dimensional movement size of the tongue and jaw, as quantified by the AWS, was reflective of self-reported speech problems as measured by the Speech Handicap Index (SHI). Insight into whether variables such as the AWS are reflective of individual experiences may contribute to the variables' clinical importance, as minimizing an individual's self-reported problems during their daily communication is a key goal of speech therapy. While no formal relationship between the AWS and the SHI has been established in previous work, one would predict that higher AWS would be related to lower SHI scores, as a larger AWS has been linked to increased speech intelligibility (Kearney et al., 2017; Lee & Bell, 2018).

Method

Participants

Speech data were taken from the Oral Cancer EMA corpus constructed by Halpern et al. (2022). The corpus contains kinematic and parallel acoustic recordings from 12 Dutch ITOC and eight control speakers. Speakers were recruited for the speech corpus if they (a) were 18 years or older, (b) were native speakers of Dutch, (c) did not have a history of neurological or speech disorders (e.g., a stroke or a stutter), (d) did not report on having any depression-related symptoms, and (e) did not have nonremovable metal (other than medical-grade titanium) in or around the head (e.g., a pacemaker or deep-brain stimulation system). ITOC additionally had to meet the criteria of having been (f) surgically treated for OSCC and (g) treated at least a year prior to the data collection taking place. ITOC were treated for tumors that were staged between T1 (smallest) and T4 (largest). All ITOC were informed of the current project by their treating clinician during regular checkups between October 2021 and April 2022.

The present study examined the kinematic data from nine ITOC (five male, four female) as well as eight age-matched controls (five male, three female) from the Halpern et al. (2022) corpus. The data of three ITOC were excluded in the current investigation as no relevant kinematic data were collected from these participants due to tongue sensors not adhering well enough for data collection. The nine ITOC received treatment either for a tumor located in the jaw (n = 5) or on the tongue (n = 4). The posttreatment time ranged from 1.2 to 11.4 years (M = 5.1 years, SD = 3.7 years). The age range of the ITOC was from 47 to 75 years (M = 61.6 years, SD = 9.4 years) and was comparable to that of the controls, with an age ranging from 56 to 77 years (M = 60.9 years, SD = 7.1 years). All but one individual treated for jaw tumors received reconstruction with a fibular osteocutaneous free flap (FOFL), wherein bone from the fibula is used to reconstruct the mandible (Hidalgo, 1989). One individual received a reconstruction with a pectoralis major myocutaneous flap (PMMF; Ariyan, 1979), which consists of using the pectoralis major muscle in combination with a titanium plate. Of the four individuals treated for tongue tumors, two received a radial forearm free flap (RFFF) reconstruction. The RFFF technique uses skin and blood vessels from the forearm to reconstruct the tongue (Yang et al., 1997). One individual treated for a tongue tumor received an anterior lateral thigh flap, which is similar to the RFFF but uses the inner thigh as the donor site instead (Song et al., 1984). Finally, one individual treated for a tongue tumor received a split skin graft, wherein a piece of skin is used to cover up the wound, as it cannot be closed locally. Details of the ITOC are provided in Table 1. All participants gave their written informed consent before participation, and the study's protocol was approved by the Medical Ethical Review Board of the University Medical Center in Groningen (NL76137.042.20).

Table 1. Demographics and clinical details of the individuals who received treatment for oral squamous cell carcinoma.

Speaker ID	Sex	Age	T-stage	SHI	Side	Location	Procedure type	Reconstruction	FOM	PORT	CRT
NKI02	M	68	T2	24	Both	Mandible	Continuity resection	FOFL	X	—	X
NKI04	M	47	T4	29	Right	Tongue	Hemiglossectomy	ALTF	X	—	X
NKI05	F	54	T4	6	Right	Mandible	Hemimandibulectomy	FOFL	X	X	—
NKI06	F	57	T3	13	Left	Tongue	Hemiglossectomy	RFFF	X	—	—
NKI11	F	56	T4	31	Both	Mandible	Continuity resection	FOFL and PMMF	X	X	—
NKI15	M	69	—	18	Both	Mandible	Total mandibulectomy	FOFL	X	—	—
NKI16	F	71	T4	9	Right	Mandible: RTA	Hemimandibulectomy	FOFL	X	—	X
NKI17	M	57	T1–T2[a]	26	Both[b]	Tongue, maxilla, cheek, oropharynx, mandible	Hemiglossectomy	RFFF	X	—	X
NKI18	M	75	T2	6	Left	Anterior tongue	Partial glossectomy	SSG	X	X	—

Note. T-stage = tumor stage; SHI = Speech Handicap Index; FOM = floor of mouth resection; PORT = postoperative radiation therapy; CRT = chemoradiation therapy; FOFL = fibular osteocutaneous free flap; ALTF = anterior lateral thigh flap; RFFF = radial forearm free flap; PMMF = pectoralis major myocutaneous flap; RTA = retromolar trigone area; SSG = split skin graft.

[a]Individual had over seven procedures for T1 or T2 malignancies. [b]Hemiglossectomy and maxillary resection were performed on the right side. Oropharyngeal tumors were removed on both sides.

Materials

Speech stimuli were taken from a pool of 226 sentences that were collected in the larger project and recorded in several blocks of six to 17 sentences each (M = 13.4 sentences; Halpern et al., 2022). To determine which block of sentences would yield the highest[1] AWS, the AWS for all sentences for two control participants was computed and averaged over each individual recording block. As sensors were checked between blocks, one block was chosen in order to control for within-speaker sensor placement differences that could occur if sensors had to be reattached in between blocks (as a sensor might not be reattached at the same place). Visual inspection of the data showed that two blocks yielded the highest AWS: a block containing sentences from The North Wind and the Sun passage and a block of sentences from newspapers. To promote reproducibility and potential cross-linguistic comparisons, The North Wind and the Sun passage was selected (see the Appendix for the Dutch passage aligned with the English version by Roach, 2004). As this passage was produced during one of the earliest blocks by the participants, this also increased the likelihood that all sensor data would still be available and that possible fatigue effects would be minimized. The passage was recorded in eight separate sentences that ranged from eight to 26 words (M = 14.6 words, SD = 5.5).

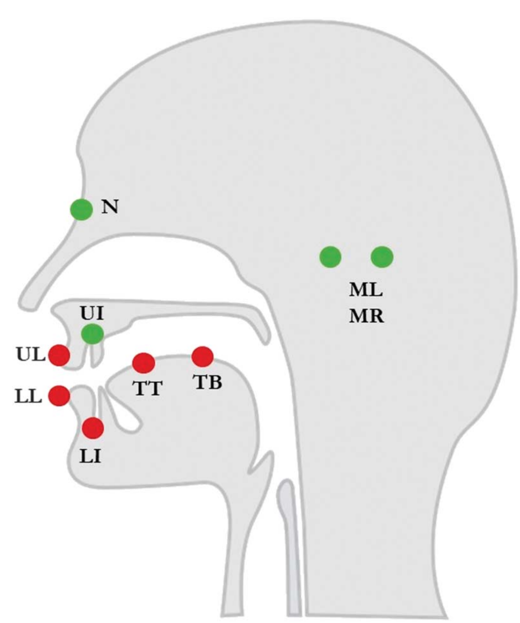
Self-reported speech outcomes in ITOC were measured using the 15-item SHI (Van den Steen et al., 2011). The version by Van den Steen et al. (2011) is a validated shortened translation of the French version of the SHI (Fichaux-Bourin et al., 2009) and contains a subset of questions from the original 30-item questionnaire for oral and oropharyngeal cancer patients by Rinkel et al. (2008). ITOC responded to the 15 items using a 5-point Likert scale ranging from *never* (0 points) to *always* (4 points). The resulting scores on the SHI range from 0 to 60, with higher scores indicating higher levels of self-reported speech problems. Mendoza Ramos et al. (2021) report the following categorization criteria: < 14, no impact; 14–22, light; 23–31, moderate; and > 31, severe impact. Van den Steen et al. (2011) found a mean SHI score of 5 for control speakers, whereas the mean score in our ITOC sample was 17.9 (SD = 9.4).

Procedure

After written informed consent was provided, the ITOC filled in the SHI. Next, all participants completed speech recordings in a sound-attenuated room of the mobile laboratory SPRAAKLAB from the Faculty of Arts, University of Groningen. Participants completed the speech tasks when the mobile laboratory was parked either at the University of Groningen or at the participant's home (in case traveling to the university was found to be difficult). Articulatory–kinematic data were collected using the Vox-EMA articulograph with a sampling rate of 400 Hz (Northern Digital Inc., 2019; Rebernik, Jacobi, Tiede, & Wieling, 2021). Acoustic data were simultaneously recorded using a Sennheiser ME66 shotgun microphone with a sampling frequency of 22050 Hz placed about 20 cm from the participant's mouth.

To collect articulatory–kinematic data, the data collection procedure specified by Rebernik, Jacobi, Jonkers, et al. (2021) was used. Five measurement sensors were attached to the articulators: one sensor on the lower incisor (jaw) to measure jaw movement, two sensors on the tongue (one approximately 1 cm from the anatomical tip [TT] and one sensor [TB] near the /k/ constriction), and two sensors on the lips (upper and lower vermilion border). Moreover, four reference sensors were used in order to correct for head movements: one on each mastoid (left and right), one on the nasal bridge, and one on the upper incisor. Figure 1 shows the placement of the sensors. A biteplate recording was made for each speaker using a protractor with three sensors attached and subsequently

Figure 1. Sensor placement. Red dots indicate measurement sensors; green dots indicate reference sensors. N = nasal bridge; ML = mastoid left; MR = mastoid right; UI = upper incisor; LI = lower incisor/jaw; TT = tongue tip; TB = tongue back; UL = upper lip; LL = lower lip.



[1]The block with the highest AWS elicits the largest range of articulatory movements and is therefore most representative of the potential maximum range of articulatory movements.

used to rotate the head-corrected data to a shared coordinate system. A palate trace was made using a spare sensor taped to the participant's thumb.

Data Preprocessing and Variable Construction

All data were head-corrected and rotated in MATLAB Version 2020a (MathWorks, 2020). All recordings were manually checked, and leading and trailing silences or audible swallowing were removed from the recording (Lee & Bell, 2018). When the participant had to reread the sentence due to a reading error, only the correctly uttered restart was used. The AWS was calculated in mm^2 for each utterance using the *convhull()* function in MATLAB for the TT, TB, and jaw sensors. Data from the TT and TB sensors were decoupled from the jaw by subtracting the jaw movement in order to assess the differences more directly. The AP-ROM and SI-ROM were calculated in mm by determining the difference between the maximum and minimum values from the X (anteroposterior) and Z (superior–inferior) coordinates for each utterance. For all three outcome measures (AWS, AP-ROM, and SI-ROM), a total of 136 data points (17 speakers × 8 pronounced sentences) were obtained for all three measurement sensors. One data point for the jaw AWS was removed due to dubious sensor tracking, yielding a total of 135 data points.

Analyses

After extracting the AWS and ROM values, data were processed in R Version 4.1.3 (R Core Team, 2020) using the tidyverse package Version 1.3.2 (Wickham et al., 2019). The resulting data set was analyzed using linear mixed-effects regression using the lme4 package Version 1.1.30 and the LmerTest package Version 3.1.3 (Bates et al., 2015; Kuznetsova et al., 2017). All outcome variables (AWS, AP-ROM, and SI-ROM) were z-transformed to assess the relative magnitude of the effects. The 95% confidence intervals were computed using the boot package Version 1.3.28 by means of a bootstrap analysis with 1,000 simulations (Canty & Ripley, 2022). The α level was set at .05.

For our first research question, in which we assessed whether two-dimensional (AWS) and one-dimensional (AP-ROM and SI-ROM) movement size was smaller in ITOC compared to control speakers, we fitted a linear mixed-effects model with the z-scored movement size as the outcome variable with sex (male–female) and the sensor plus measurement combination (e.g., TT AWS or TB AP-ROM) in interaction with group (ITOC–control) as the independent variables. We included random intercepts for sentence and speaker and a random slope for group per sentence.

For our second research question, which assessed whether the size of the two-dimensional AWS was related to the primary treatment site, we fitted a linear mixed-effects model with the z-scored movement size for the AWS with sex (male–female) and sensor (jaw, TT, and TB) in interaction with subgroup (control, tongue, jaw) as the independent variables. We included random intercepts for sentence and speaker. A random slope for group per sentence led to a singular fit and was therefore not included.

Finally, for our third research question, which assessed whether the size of the two-dimensional AWS was related to the self-reported speech problems as measured by the SHI, we ran a Spearman's rank correlation test between the mean-centered SHI score and the mean of the AWS per sensor across all eight sentences, resulting in three tests. Due to our small clinical sample (n = 9), we used a nonparametric test and focused on effect size rather than significance. Only correlations that were at least of moderate strength ($|r_s| \geq .4$) were interpreted.

Hypothesis testing for Questions 1 and 2 was followed by an exploratory analysis assessing the influence of speaker age. Model selection was done using the *anova()* function, where a p value below .05 was used to identify that the more complex model was preferred. After the optimal model was selected, model criticism was employed by refitting the final models to the original data set but removing data points whose residuals were at least 2 *SD*s away from the fitted value. In order to not report results that were driven by outliers, we used this trimmed data set only if the outliers drove the absence or presence of significant effects (Baayen, 2008, Chapter 6.2.3). Our final models adhered to the assumptions of normality and homoscedasticity. All data visualizations were made with the ggplot2 package Version 3.3.6 (Wickham, 2016).

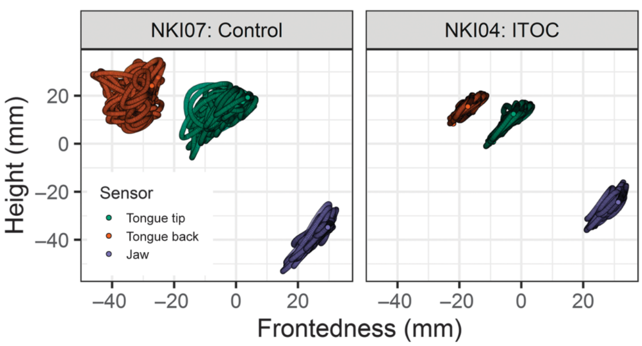
Results

Group Comparison

A sample AWS for all three sensors from one control speaker and an individual treated for tongue cancer is provided in Figure 2. Figures 3 and 4 visualize the results for the AWS, AP-ROM, and SI-ROM measures on both the individual level and the group level. The model output is summarized in Table 2, whereas the individual- and group-level descriptives are provided in Table 3.

While controlling for sex, the AWS was significantly smaller for ITOC compared to control speakers for the TT (β = −0.84 *SD*, t = −5.6, p < .001) and the TB

Figure 2. Two examples of the articulatory working space (AWS) of the tongue tip (TT; in green), tongue back (TB; in orange), and jaw (in purple). Higher values in the anteroposterior dimension correspond to a more fronted sensor position, whereas higher values in the superior–inferior dimension correspond to a higher sensor position. The left side of the plot shows the values for NKI07, a male control speaker, who had one of the highest AWS values for all three sensors. The right side shows the values for NKI04, a male treated for oral cancer (ITOC) who had very restricted tongue movements (one of the lowest AWS for TT and TB). ITOC = individual treated for oral squamous cell carcinoma.



(β = −0.59 *SD*, $t = -3.9$, $p < .001$), but not for the jaw (β = 0.03 *SD*, $t = 0.23$, $p = .82$). The AP-ROM was significantly smaller for ITOC compared to control speakers for both the TT sensor (β = −0.58 *SD*, $t = -3.8$, $p < .001$) and the TB sensor (β = −0.38 *SD*, $t = -2.6$, $p = .01$), but not for the jaw sensor (β = −0.26 *SD*, $t = -0.91$, $p = .37$). The SI-ROM was significantly smaller for ITOC compared to control speakers for both the TT sensor (β = −0.88 *SD*, $t = -5.9$, $p < .001$) and the TB sensor (β = −0.69 *SD*, $t = -4.6$, $p < .001$), but not for the jaw sensor (β = 0.17 *SD*, $t = 1.1$, $p = .26$). The subsequent exploratory analysis investigating the effect of age did not yield a significant result, $\chi^2 = 1.83(1)$, $p = .18$.

Effect of Treatment Site on Kinematic Measures

While controlling for sex, the AWS of the TT sensor was significantly smaller compared to control speakers for both individuals treated for tongue cancer (β = −0.91 *SD*, $t = -6.6$, $p < .001$) and jaw cancer (β = −1.0 *SD*, $t = -7.8$, $p < .001$). Similarly, the AWS of the TB sensor was significantly smaller compared to control speakers for both individuals treated for tongue cancer (β = −0.75 *SD*, $t = -6.4$, $p < .001$) and jaw cancer (β = −0.84 *SD*, $t = -6.4$, $p < .001$). No significant differences were found for the AWS of the jaw for individuals treated for tongue cancer (β = −0.26 *SD*, $t = -0.66$, $p = .51$) or jaw cancer (β = −0.07 *SD*, $t = -0.19$, $p = .85$; see Table 4). The subsequent exploratory analysis investigating the effect of age did not yield a significant result, $\chi^2 = 2.21(1)$, $p = .14$.

Relation Between AWS and SHI

The mean SHI score for ITOC was 17.9 (*SD* = 9.4), and all ITOC rated their speech problems as either light or mild following treatment. Figure 5 shows the mean SHI score and the mean of the TT, TB, and jaw AWS values for each individual participant together with the trend line of the ITOC group as a whole. There was no significant correlation between the SHI score for both the TT AWS ($r_s = -.29$, $p = .4$) and the TB AWS ($r_s = -.4$, $p = .3$). However, the correlation between SHI and TB AWS was of moderate effect size. The Spearman's rank correlation test between the SHI score and the AWS of the jaw resulted likewise in a nonsignificant correlation, but here the effect size was moderate as well ($r_s = .49$, $p = .2$).

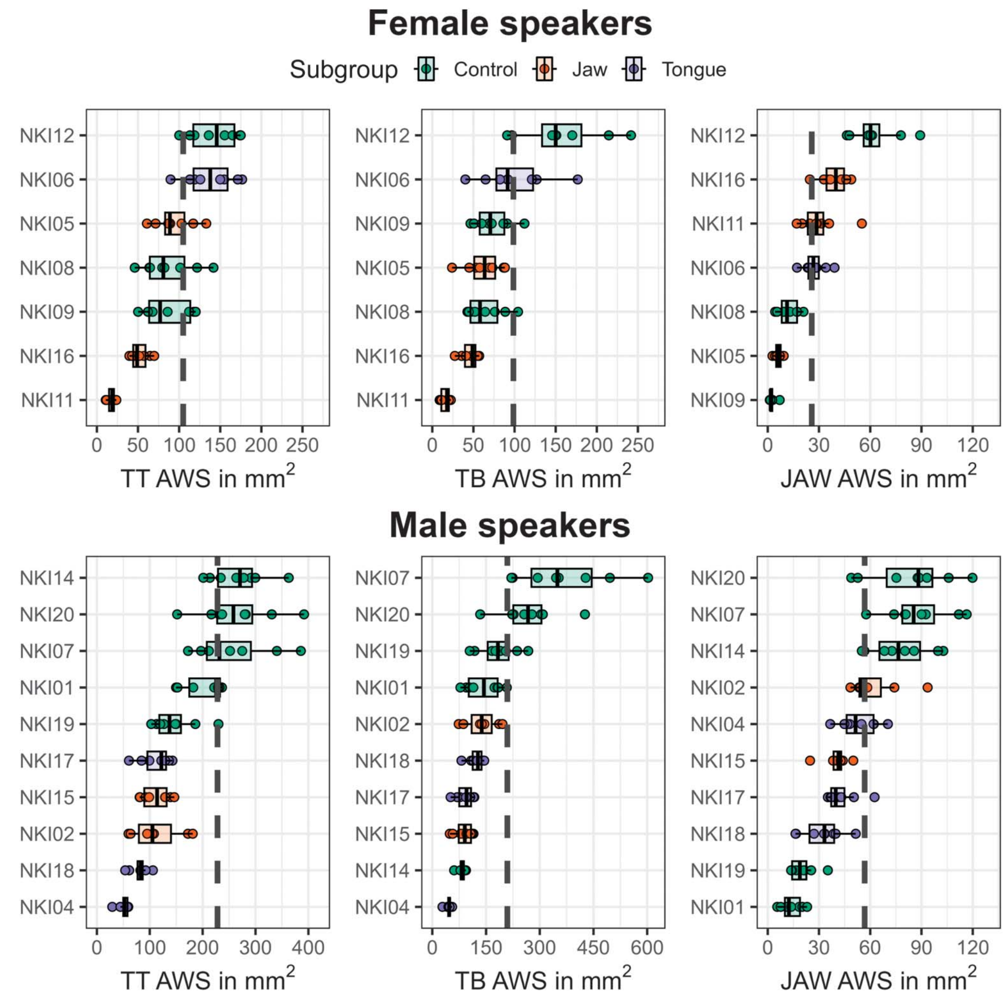
Discussion

Group Comparison

The purpose of this study was to assess the articulatory function of nine ITOC compared to eight control speakers by using both one-dimensional (AP-ROM and SI-ROM) and two-dimensional (AWS) sentence-level articulatory–kinematic data. Kinematic data were acquired using EMA during a reading of The North Wind and the Sun passage.

Our first aim was to assess whether the AWS, AP-ROM, and SI-ROM would be reduced in ITOC when compared to control speakers. Our hypothesis was that ITOC would have smaller AWS values compared to the control speakers as OSCC treatment may limit the mobility of the articulators (Bressmann et al., 2004; Chepeha et al., 2016; de Groot et al., 2020; Kappert et al., 2019; Speksnijder et al., 2011). In line with our hypothesis and previous work, we found significant differences for both the TT and TB AWS. Bressmann et al. (2004) previously found that tongue mobility, as measured on a 3-point Likert scale, correlated positively with consonant intelligibility in individuals who received a partial glossectomy. The authors further noted that different motor demands may exist between speech and nonspeech tasks. Whereas nonspeech tasks measure maximal movements in terms of anatomical capability, speech tasks require more fine-grained motor movement. Our results show that ITOC have reduced tongue mobility during a reading task that requires more fine-grained motor commands as compared to control speakers. However, no direct comparisons between speech and nonspeech tasks were made in the present study. Future research should assess to what extent kinematic measures of mobility during speech (e.g., AWS) and clinically implemented nonspeech tasks (e.g.,

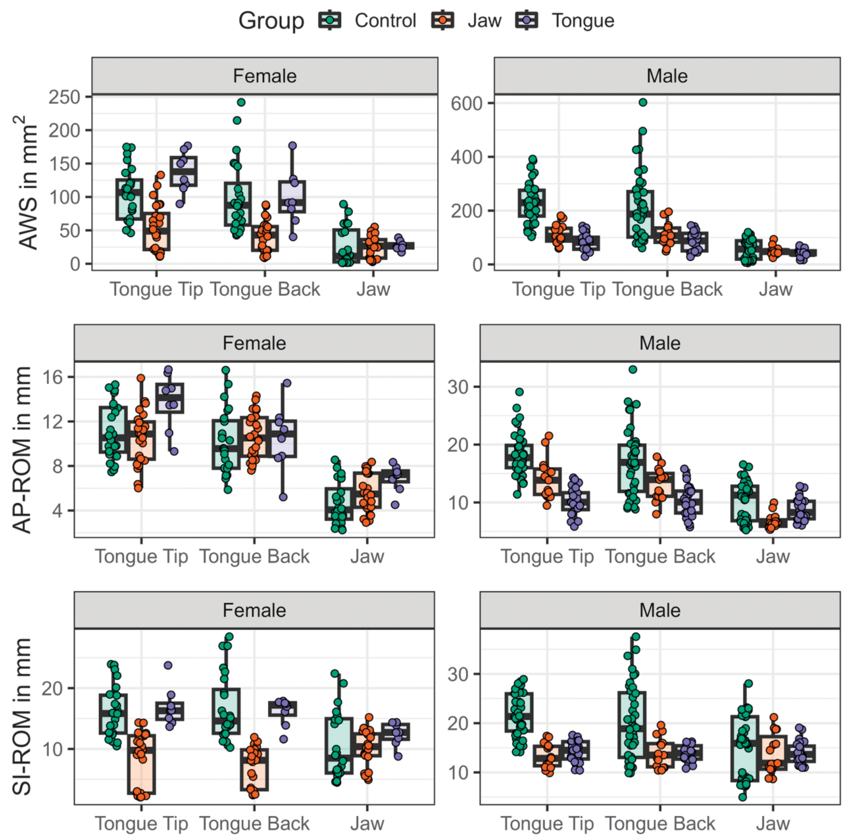
Figure 3. Articulatory working space (AWS) measures (in mm^2) for control speakers (green) and individuals treated for a tongue tumor (purple) or a jaw tumor (orange) for both female (upper plot) and male (lower plot) speakers. Note that the tongue tip (TT) and tongue back (TB) AWS are plotted with different axis limits for females (0–250 mm^2) and males (0-400 and 0–600 mm^2). The gray dashed line marks the control mean, and the dots represent individual measurements of the AWS.



ruler- or Likert scale–based assessments) capture the same information, considering that EMA is not suitable to be used in clinical practice, whereas quick ruler-based assessments are.

Besides differences in two-dimensional movement (the AWS), we also found significant differences for the one-dimensional SI-ROM and AP-ROM for both the TT and TB sensors, which partly agree with our hypotheses and previous work. We hypothesized that the AP-ROM, but not the SI-ROM, would be reduced in ITOC compared to control speakers based on prior work using speech acoustics (Acher et al., 2014; de Bruijn et al., 2009; Laaksonen et al., 2011; Takatsu et al., 2017; Tienkamp et al., 2023; Zhou et al., 2013). One possible explanation for the discrepancy between our hypothesis and the results regarding reduced SI-ROM is that our hypothesis was based on acoustic studies. Speech acoustics provide only an approximation of the underlying movements of the tongue and may be affected by the compensatory use of other articulators (e.g., jaw and lips). For example, Hagedorn et al. (2022) suggests that ITOC might use the jaw in a compensatory manner in order to compensate for reduced tongue mobility. In turn, this jaw movement may result in typical first formant values, which are affected by both tongue height and jaw positioning. Though nonsignificant, the ITOC in our sample displayed larger SI-ROM of the jaw, which could be a sign of compensatory movement. Thus, compensatory mechanisms might mask effects for vertical tongue movement in articulatory–acoustic analyses, whereas limited SI-ROM of the tongue is still present in articulatory–kinematic studies where the movement of the tongue and jaw are measured separately. This highlights the importance of kinematic methods in

Figure 4. Articulatory working space (AWS; in mm^2), anteroposterior range of motion (AP-ROM; in mm), and superior–inferior range of motion (SI-ROM; in mm) for control speakers (green) and individuals treated for a jaw tumor (orange) or a tongue tumor (purple) stratified by sex (female–male). Note that the axis limits are different for males and females. Dots represent individual measurements of the outcome variables.



two ways. First, it helps to establish successful compensatory strategies that may be taught to other ITOC as well. Second, it provides direct evidence of impacted movements that may be addressed in therapeutic interventions. Recently, Blyth et al. (2023) documented that SLPs focus primarily on compensation rather than active rehabilitation. Importantly, the SLPs also acknowledged the scarcity of relevant evidence as a barrier to practice. Thus, kinematic data are of crucial importance to guide clinical rehabilitation. Overall, ITOC showed reduced tongue

Table 2. Summary of kinematic movement variables for the group comparison, model: *z*-score ~ combination × group + sex + (1 + group | sentence) + (1 | speaker).

Parameter	Estimate (95% CI)	*SE*	*df*	*t*	*p*
Jaw AWS × ITOC	0.03 [−0.25, 0.33]	0.15	1168	0.23	.82
TT AWS × ITOC	−0.84 [−1.13, −0.56]	0.15	1168	−5.6	< .001
TB AWS × ITOC	−0.59 [−0.88, −0.3]	0.15	1168	−3.9	< .001
Jaw AP-ROM × ITOC	−0.26 [−0.85, 0.29]	0.29	19	−0.91	.37
TT AP-ROM × ITOC	−0.58 [−0.88, −0.28]	0.15	1168	−3.8	< .001
TB AP-ROM × ITOC	−0.38 [−0.69, −0.09]	0.15	1168	−2.56	.01
Jaw SI-ROM × ITOC	0.17 [−0.14, 0.46]	0.15	1168	1.13	.26
TT SI-ROM × ITOC	−0.88 [−1.19, −0.58]	0.15	1168	−5.88	< .001
TB SI-ROM × ITOC	−0.69 [−1.0, −0.39]	0.15	1168	−4.59	< .001

Note. CI = confidence interval; *SE* = standard error; *df* = degrees of freedom; AWS = articulatory working space; ITOC = individuals treated for oral squamous cell carcinoma; TT = tongue tip; TB = tongue back; AP-ROM = anteroposterior range of motion; SI-ROM = superior–inferior range of motion.

Table 3. Mean and standard deviation for each outcome measure for each individual who received treatment for oral squamous cell carcinoma and the control and ITOC groups as a whole.

Participant	Subgroup	SHI	Articulatory working space in mm^2 (*SD*)			Anteroposterior ROM in mm (*SD*)			Superior–inferior ROM in mm (*SD*)		
			TT	TB	Jaw	TT	TB	Jaw	TT	TB	Jaw
Females											
Controls	—	—	104.9 (39.2)	98.6 (54.3)	25.8 (28.3)	11.1 (2.5)	10.2 (2.8)	4.6 (1.9)	16.1 (4.2)	16.7 (5.5)	10.5 (5.5)
ITOC	—	14.8 (9.9)	75.2 (49.5)	55.8 (38.3)	25.7 (14.4)	11.3 (2.8)	10.6 (2.3)	5.9 (1.7)	10.5 (5.6)	9.3 (5.0)	10.8 (2.9)
NKI05	Jaw	6	93.7 (23.4)	62.1 (21.8)	6.2 (2.3)	12.4 (1.9)	10.6 (2.0)	4.0 (1.2)	12.5 (1.4)	9.3 (2.5)	7.0 (2.1)
NKI11	Jaw	31	17.7 (4.6)	16.1 (5.2)	30.2 (11.8)	10.7 (2.0)	11.5 (2.0)	5.2 (1.0)	2.5 (0.3)	2.9 (0.4)	12.4 (1.5)
NKI16	Jaw	9	52.1 (10.9)	45.7 (10.2)	38.9 (8.1)	8.6 (2.0)	9.9 (2.1)	7.6 (0.6)	10.2 (1.4)	8.9 (0.8)	11.2 (1.9)
NKI06	Tongue	13	137.5 (30.5)	99.4 (42.0)	27.4 (6.8)	13.8 (2.6)	10.5 (3.0)	6.9 (1.2)	16.9 (3.2)	16.1 (2.3)	12.5 (1.9)
Males											
Controls	—	—	228.1 (72.5)	290.5 (126.2)	56.7 (36.6)	18.3 (3.7)	17.1 (5.8)	10.2 (3.4)	21.8 (4.4)	20.0 (7.7)	14.9 (6.3)
ITOC	—	20.5 (8.3)	93.1 (36.6)	95.7 (40.6)	45.7 (15.0)	11.7 (4.0)	11.4 (3.1)	8.0 (2.0)	13.9 (2.3)	14.0 (2.2)	13.8 (3.1)
NKI02	Jaw	24	119.1 (48.5)	141.6 (44.3)	80.7 (54.1)	16.0 (3.7)	14.0 (2.6)	8.4 (2.6)	12.6 (2.3)	15.7 (2.5)	18.7 (3.6)
NKI15	Jaw	18	112.2 (26.7)	87.4 (24.6)	40.2 (7.2)	12.5 (2.1)	12.4 (2.6)	6.0 (0.5)	14.2 (2.6)	12.7 (3.0)	10.5 (1.2)
NKI04	Tongue	29	50.7 (10.0)	45.6 (8.0)	53.2 (11.1)	8.2 (1.4)	8.3 (1.5)	10.4 (1.5)	15.0 (1.7)	14.1 (1.4)	15.6 (2.7)
NKI17	Tongue	26	112.1 (28.3)	90.8 (23.2)	43.11 (9.2)	11.6 (2.6)	9.3 (2.0)	7.5 (1.3)	14.8 (1.7)	14.2 (1.5)	12.8 (1.3)
NKI18	Tongue	6	80.7 (16.6)	122.6 (20.9)	32.0 (12.0)	10.7 (1.3)	13.2 (1.5)	8.6 (2.0)	13.2 (2.6)	13.4 (2.0)	13.2 (1.9)

Note. Results are stratified by sex (females in the upper half and males in the lower half of the table). ITOC = individuals treated for oral squamous cell carcinoma; SHI = speech handicap index (ranges from 0 to 60); ROM = range of motion; TT = tongue tip; TB = tongue back.

Table 4. Summary of kinematic movement variables for the subgroup comparison, model: *z* score ~ sensor × subgroup + sex + (1 | sentence) + (1 | speaker).

Parameter	Estimate (95% CI)	*SE*	*df*	*t*	*p*
TT AWS × Jaw ITOC	−1.0 [−1.24, −0.61]	0.13	351.1	−7.77	< .001
TT AWS × Tongue ITOC	−0.91 [−1.18, −0.49]	0.14	351.1	−6.64	< .001
TB AWS × Jaw ITOC	−0.84 [−0.98, −0.33]	0.13	351.1	−6.43	< .001
TB AWS × Tongue ITOC	−0.75 [−0.9, −0.25]	0.14	351.1	−5.41	< .001
Jaw AWS × Jaw ITOC	−0.07 [−0.85, 0.56]	0.37	14.1	−0.19	.851
Jaw AWS × Tongue ITOC	−0.26 [−1.13, 0.45]	0.39	14.1	−0.66	.51

Note. CI = confidence interval; *SE* = standard error; *df* = degrees of freedom; TT = tongue tip; AWS = articulatory working space; ITOC = individuals treated for oral squamous cell carcinoma; TB = tongue back.

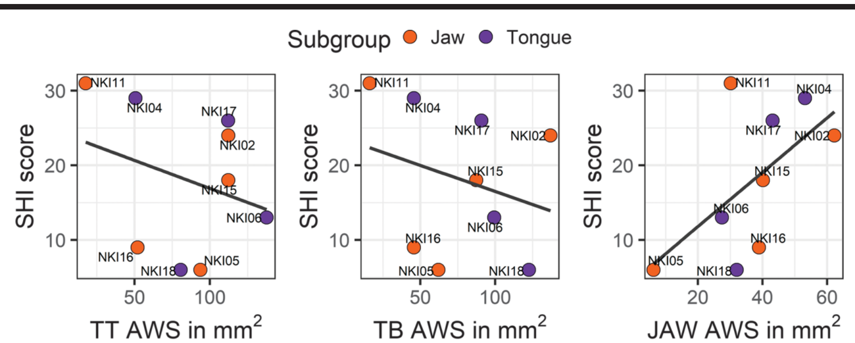
movement in both the anteroposterior and superior–inferior dimension, which could be due to either scar tissue and/ or tissue loss from the surgical procedure or fibrosis due to postoperative radiation applied to the primary treatment site.

At the group level, no significant differences were found for the AWS, AP-ROM, and SI-ROM of the jaw, which could be due to the bimodal distribution of jaw movement in the control group. For example, there are control speakers who use the jaw to a considerable extent (e.g., NKI12, NKI07, NKI20), whereas other control speakers do not use the jaw as much (e.g., NKI01, NKI09; see Figure 3). In our sample, both male and female control speakers were at the extreme ends of both sides of the jaw movement continuum, whereas the ITOC were in the middle of this continuum (see Figure 3). On the one hand, this bimodal distribution may mask effects of reduced jaw movements in the data. On the other hand, the data suggest that only very minimal jaw movement is needed to produce typical speech. Another reason for the absence of a group-level effect is that both individuals treated for tongue and jaw tumors were included in the group analysis. We return to this issue below.

Effect of Treatment Site on Kinematic Measures

The second aim of our study was to assess whether the aforementioned differences in AWS size could, in part, be explained by the primary treatment site. We hypothesized that tongue AWS would be reduced primarily in individuals treated for tongue tumors and when jaw tumors were in proximity to the tongue (i.e., when parts of the floor of the mouth had to be resected as well; Chepeha et al., 2016; de Groot et al., 2020; Kappert et al., 2019; van Dijk et al., 2016). Additionally, we hypothesized that the jaw AWS would be reduced in individuals treated for tumors in the jaw. Our results partly support this hypothesis. We found that TT and TB AWS were significantly smaller for both tumors on the tongue and jaw compared to control speakers. The results are in line with our hypothesis as all individuals treated for jaw tumors received a resection of the floor of the mouth. Moreover, a possible effect of sublocation was found. NKI18 (male, T2, partial glossectomy, split skin graft) was treated for a tumor on the anterior part of the tongue. Both the AWS (80.7 mm^2) and AP-ROM (10.7 mm) were lower for the TT, the affected region, compared to the AWS

Figure 5. Scatter plot with the Speech Handicap Index (SHI) score and the mean tongue tip (TT), tongue back (TB), and jaw articulatory working space (AWS; in mm^2) measures. Individuals treated for a jaw tumor are plotted in orange circles, and those treated for a tongue tumor are plotted in purple. The speaker ID is plotted next to the data point. The gray line plots the trend for the individuals treated for oral squamous cell carcinoma (ITOC) group as a whole.



(122.6 mm^2) and AP-ROM (13.4 mm) of the TB, the unaffected region. This suggests that the primary treatment site is affected more than surrounding structures in this individual in terms of mobility during speech.

No significant differences were found between individuals treated for jaw tumors and control speakers for the jaw AWS. While the bimodal distribution of the control group data might serve as an explanation, it could also be the case that treatment for jaw tumors does not severely impact the mobility of the jaw, as long as treatment does not intervene with the structural parts associated with jaw movement (e.g., the temporomandibular joint and masticatory muscles). For example, reconstructions using an FOFL do not affect these structures directly. However, NKI17 and NKI16 had tumors removed in the retromolar trigone area, which could affect jaw mobility, but also had jaw AWS values that fell within typical ranges. It should be noted, though, that only posttreatment data were collected in our sample, meaning we cannot assess whether NKI16 and NKI17 had larger jaw AWS pretreatment that was subsequently reduced due to the surgical intervention. Lastly, though radiation therapy applied to primary tumor sites could affect the jaw's mobility by inducing muscle stiffness, we did not find evidence for reduced mobility during speech production.

Relation Between AWS and SHI

The third aim of our study was to assess the degree to which the AWS was reflective of self-reported speech problems. Based on the link between the AWS and speech intelligibility, we predicted that higher AWS would be related to lower SHI scores (i.e., lower levels of self-reported speech problems). Our results show no significant correlations between the AWS for all three sensors and the SHI score. However, this is not surprising considering our small clinical sample of nine speakers. Moreover, the SHI scores of the included speakers indicated no to moderate impact as scores were 31 or less, with a mean score of 17.9 (range: 6–31). This could serve as an additional explanation as to why no significant correlations were found. Therefore, we focus more on effect size rather than significance, and only correlations that were at least of moderate strength ($|r_s| \geq .4$) are interpreted here. In this regard, there was a negative correlation of moderate strength for the TB ($r_s = -.4$), meaning that higher SHI scores were associated with lower TB AWS, which is in line with our prediction. Qualitatively, the two speakers with the highest SHI score, NKI11 (female, T4, continuity resection, PMMF, SHI = 31) and NKI04 (male, T4, hemiglossectomy, anterolateral thigh flap, SHI = 29), also displayed the lowest TB AWS. The opposite pattern was found for the jaw AWS where a moderate positive correlation ($r_s = .49$) indicated that higher SHI scores were associated with higher jaw AWS. This finding may be in line with the possibility that our speakers use the jaw in a compensatory manner. Though a larger sample size is needed in order to draw stronger conclusions concerning the relationship between the AWS and self-reported speech outcomes, our results tentatively suggest that the AWS of the affected articulator (i.e., the tongue) might be related to the SHI.

Limitations

Though the present study presented novel results regarding sentence-level kinematic movement following surgical treatment for OSCC, the study was not without limitations. The present study did not collect any information regarding the ROM in nonspeech tasks (i.e., ruler-based or Likert scale–based assessments), which limits our ability to interpret to what extent one- or two-dimensional movement data during speech and nonspeech tasks relate to each other. This would be a fruitful direction for future research, as establishing a connection between speech and nonspeech ROM could validate the use of ruler-based ROM measurements or Likert scale–based assessments to approximate tongue and jaw movement during speech.

Another limitation of the study is that we did not collect any information regarding the intelligibility of the speakers, so comparisons between the size of the AWS and intelligibility or clarity of speech cannot be assessed. Previous studies have established a positive relationship between AWS size and intelligibility or speech clarity in dysarthric populations, such that higher AWS values corresponded to higher intelligibility (Kearney et al., 2017; Lee & Bell, 2018). These relationships were found for individuals with neurodegenerative diseases that affect the speech motor system rather than purely physiological changes as with ITOC. However, it should be noted that a small AWS need not immediately result in reduced intelligibility. For example, intelligibility may be maintained for speakers having a small AWS if phonemes are produced with maximal distinction within this smaller space (Lee & Bell, 2018; Weismer, 2013). This would require considerably precise motor coordination. One way this could be quantified with EMA is to use an articulatory consonant distinctiveness approach in which kinematic trajectories from multiple sensors (e.g., tongue, lip, and jaw) are combined in order to calculate a composite shape of a given phoneme (Teplansky et al., 2023; Wang et al., 2013). The distance between these phoneme shapes serves as an index of how distinct the consonants are realized kinematically. Moreover, an indication of the movement size might be obtained if multidimensional scaling is applied to the obtained distances. Teplansky et al. (2023)

have shown that a smaller consonant space was associated with greater clinical severity ratings in speakers with amyotrophic lateral sclerosis. Thus, future research could address the link between AWS and intelligibility or clarity of speech in ITOC using this approach, considering that this population might have difficulty in preserving precise motor coordination and maximal phoneme distinction within a smaller AWS.

Due to the anatomic alterations introduced by oral cancer treatment, the sensor placement differed slightly per individual. This may have introduced additional variability in our sample of ITOC. Even though the distance between the TT and TB sensor was similar for ITOC and controls in most cases, the distance was reduced for some in the case of major tissue loss. This means that the differential contributions of the TT and TB in these individuals are less easily assessed. However, this is inevitable in this population due to the treatment methods.

Finally, due to the relatively small clinical sample size of nine speakers, combined with the experience that oral cancer treatment outcomes are highly variable, the results of the study are limited in their generalizability. A general shortcoming of studies using the AWS is that there is no satisfactory way to control for individual differences in terms of oral cavity size. Oral cavity size might serve as an explanation for the considerable individual variation as larger movements are necessary in larger oral cavities. We tried to account for these differences to a certain extent by controlling for sex in our analyses (i.e., males have, on average, larger oral cavities than females).

Conclusions

To the best of our knowledge, the study presented the first report of sentence-level articulatory–kinematic measures of one- and two-dimensional movements in ITOC using EMA. Despite large individual variation, ITOC had a significantly smaller AWS, AP-ROM, and SI-ROM for both the TT and TB compared to control speakers. Both treatment for tongue and jaw tumors resulted in reduced tongue mobility during speech, replicating studies that employed nonspeech tasks. No significant differences were found between the groups in terms of jaw movement. A larger sample size is needed to corroborate the moderate, but nonsignificant, correlations between the AWS and self-reported speech problems. Overall, our findings specify specific kinematic changes in ITOC and should be further explored in future studies as to how kinematic changes pertain to acoustic and perceptual changes. Kinematic data may be especially informative in designing new therapeutic approaches for this population as they directly relate to articulatory–kinematic changes induced by surgical treatment.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported by a University of Groningen Center for Language and Cognition PhD grant awarded to Thomas B. Tienkamp. This work was also supported by a research grant from Atos Medical (Hörby, Sweden), which contributes to the existing infrastructure for quality of life research, awarded to the Department of Head and Neck Oncology and Surgery of the Netherlands Cancer Institute. The authors are most grateful to all speakers who participated in this study.

References

Acher, A., Perrier, P., Savariaux, C., & Fougeron, C. (2014). Speech production after glossectomy: Methodological aspects. *Clinical Linguistics & Phonetics, 28*(4), 241–256. https://doi.org/10.3109/02699206.2013.802015

Ariyan, S. (1979). The pectoralis major myocutaneous flap: A versatile flap for reconstruction in the head and neck. *Plastic and Reconstructive Surgery, 63*(1), 73—81. https://doi.org/10.1097/00006534-197901000-00012

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press. https://doi.org/10.1017/CBO9780511801686

Bak, M., Jacobson, A. S., Buchbinder, D., & Urken, M. L. (2010). Contemporary reconstruction of the mandible. *Oral Oncology, 46*(2), 71–76. https://doi.org/10.1016/j.oraloncology.2009.11.006

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blyth, K., Stainlay, H., & McCabe, P. (2023). Speech and swallowing intervention following oral cancer treatment: A survey of speech-language pathologists in Australia and New Zealand. *International Journal of Speech-Language Pathology,* 1–11. https://doi.org/10.1080/17549507.2023.2240043

Bressmann, T., Ackloo, E., Heng, C.-L., & Irish, J. C. (2007). Quantitative three-dimensional ultrasound imaging of partially resected tongues. *Otolaryngology—Head & Neck Surgery, 136*(5), 799–805. https://doi.org/10.1016/j.otohns.2006.11.022

Bressmann, T., Sader, R., Whitehill, T. L., & Samman, N. (2004). Consonant intelligibility and tongue motility in patients with partial glossectomy. *Journal of Oral and Maxillofacial Surgery, 62*(3), 298–303. https://doi.org/10.1016/j.joms.2003.04.017

Canty, A., & Ripley, B. D. (2022). *boot: Bootstrap R (S-plus) functions*. https://CRAN.R-project.org/package=boot

Chepeha, D. B., Spector, M. E., Chinn, S. B., Casper, K. A., Chanowski, E. J. P., Moyer, J. S., Morrison, R., Carvill, E., & Lyden, T. H. (2016). Hemiglossectomy tongue reconstruction: Modeling of elevation, protrusion, and functional outcome using receiver operator characteristic curve. *Head & Neck, 38*(7), 1066–1073. https://doi.org/10.1002/hed.24417

Crystal, T. H., & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. *The Journal of the Acoustical Society of America, 72*(3), 705–716. https://doi.org/10.1121/1.388251

de Bruijn, M. J., ten Bosch, L., Kuik, D. J., Quené, H., Langendijk, J. A., Leemans, C. R., & Verdonck-de Leeuw, I. M. (2009). Objective acoustic–phonetic speech analysis in patients treated for oral or oropharyngeal cancer. *Folia Phoniatrica et Logopaedica, 61*(3), 180–187. https://doi.org/10.1159/000219953

de Groot, R. J., Merkx, M. A. W., Hamann, M. N. S., Brand, H. S., de Haan, A. F. J., Rosenberg, A. J. W. P., & Speksnijder, C. M. (2020). Tongue function and its influence on masticatory performance in patients treated for oral cancer: A five-year prospective study. *Supportive Care in Cancer, 28*(3), 1491–1501. https://doi.org/10.1007/s00520-019-04913-y

Fairbanks, G. (1960). *Voice and articulation drillbook* (Vol. 2). Harper & Row.

Fichaux-Bourin, P., Woisard, V., Grand, S., Puech, M., & Bodin, S. (2009). Validation d'un questionnaire d'auto-évaluation de la parole (parole handicap index) [Validation of a questionnaire for the automatic evaluation of speech (Speech Handicap Index)]. *Revue de laryngologie, d'otologie et de Rhinologie, 130*(1), 45–51.

Hagedorn, C., Kim, J., Sinha, U., Goldstein, L., & Narayanan, S. S. (2021). Complexity of vocal tract shaping in glossectomy patients and typical speakers: A principal component analysis. *The Journal of the Acoustical Society of America, 149*(6), 4437–4449. https://doi.org/10.1121/10.0004789

Hagedorn, C., Lu, Y., Toutios, A., Sinha, U., Goldstein, L., & Narayanan, S. (2022). Variation in compensatory strategies as a function of target constriction degree in post-glossectomy speech. *JASA Express Letters, 2*(4) Article 045205. https://doi.org/10.1121/10.0009897

Halpern, B. M., Rebernik, T., Tienkamp, T. B., van Son, R., van den Brekel, M., Wieling, M., Witjes, M., & Scharenborg, O. (2022). *Manipulation of oral cancer speech using neural articulatory synthesis*. ArXiv Preprint. https://doi.org/10.48550/arXiv.2203.17072

Hidalgo, D. A. (1989). Fibula free flap: A new method of mandible reconstruction. *Plastic and Reconstructive Surgery, 84*(1), 71–79. https://doi.org/10.1097/00006534-198907000-00014

Husaini, H., Krisciunas, G. P., Langmore, S., Mojica, J. K., Urken, M. L., Jacobson, A. S., & Lazarus, C. L. (2014). A survey of variables used by speech-language pathologists to assess function and predict functional recovery in oral cancer patients. *Dysphagia, 29*(3), 376–386. https://doi.org/10.1007/s00455-014-9520-2

Jacobi, I., van Rossum, M. A., van der Molen, L., Hilgers, F. J. M., & van den Brekel, M. W. M. (2013). Acoustic analysis of changes in articulation proficiency in patients with advanced head and neck cancer treated with chemoradiotherapy. *Annals of Otology, Rhinology & Laryngology, 122*(12), 754–762. https://doi.org/10.1177/000348941312201205

Kappert, K., van Alphen, M., Smeele, L., Balm, A., & van der Heijden, F. (2019). Quantification of tongue mobility impairment using optical tracking in patients after receiving primary surgery or chemoradiation. *PLOS ONE, 14*(8) Article e0221593. https://doi.org/10.1371/journal.pone.0221593

Kearney, E., Giles, R., Haworth, B., Faloutsos, P., Baljko, M., & Yunusova, Y. (2017). Sentence-level movements in Parkinson's disease: Loud, clear, and slow speech. *Journal of Speech, Language, and Hearing Research, 60*(12), 3426–3440. https://doi.org/10.1044/2017_JSLHR-S-17-0075

Konstantinović, V. S., & Dimić, N. D. (1998). Articulatory function and tongue mobility after surgery followed by radiotherapy for tongue and floor of the mouth cancer patients. *British Journal of Plastic Surgery, 51*(8), 589–593. https://doi.org/10.1054/bjps.1998.0202

Kreeft, A. M., van der Molen, L., Hilgers, F. J., & Balm, A. J. (2009). Speech and swallowing after surgical treatment of advanced oral and oropharyngeal carcinoma: A systematic review of the literature. *European Archives of Oto-Rhino-Laryngology, 266*(11), 1687–1698. https://doi.org/10.1007/s00405-009-1089-2

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTestPackage: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13). https://doi.org/10.18637/jss.v082.i13

Laaksonen, J.-P., Rieger, J., Harris, J., & Seikaly, H. (2011). A longitudinal acoustic study of the effects of the radial forearm free flap reconstruction on sibilants produced by tongue cancer patients. *Clinical Linguistics & Phonetics, 25*(4), 253–264. https://doi.org/10.3109/02699206.2010.525681

Lam, L., & Samman, N. (2013). Speech and swallowing following tongue cancer surgery and free flap reconstruction—A systematic review. *Oral Oncology, 49*(6), 507–524. https://doi.org/10.1016/j.oraloncology.2013.03.001

Lazarus, C. L., Husaini, H., Jacobson, A. S., Mojica, J. K., Buchbinder, D., Okay, D., & Urken, M. L. (2014). Development of a new lingual range-of-motion assessment scale: Normative data in surgically treated oral cancer patients. *Dysphagia, 29*(4), 489–499. https://doi.org/10.1007/s00455-014-9534-9

Lee, J., & Bell, M. (2018). Articulatory range of movement in individuals with dysarthria secondary to amyotrophic lateral sclerosis. *American Journal of Speech-Language Pathology, 27*(3), 996–1009. https://doi.org/10.1044/2018_AJSLP-17-0064

Lee, J., Littlejohn, M. A., & Simmons, Z. (2017). Acoustic and tongue kinematic vowel space in speakers with and without dysarthria. *International Journal of Speech-Language Pathology, 19*(2), 195–204. https://doi.org/10.1080/17549507.2016.1193899

MathWorks. (2020). *MATLAB 2020 (Version a)*. https://mathworks.com/?s_tid=gn_logo

Mefferd, A. S., & Green, J. R. (2010). Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations. *Journal of Speech, Language, and Hearing Research, 53*(5), 1206–1219. https://doi.org/10.1044/1092-4388(2010/09-0083)

Mendoza Ramos, V., Paulyn, C., Van den Steen, L., Hernandez-Diaz Huici, M. E., De Bodt, M., & Van Nuffelen, G. (2021). Effect of boost articulation therapy (BArT) on intelligibility in adults with dysarthria. *International Journal of Language & Communication Disorders, 56*(2), 271–282. https://doi.org/10.1111/1460-6984.12595

Northern Digital, Inc. (2019). *Vox-EMA system user guide*. http://support.ndigital.com

Pauloski, B. R., Logemann, J. A., Colangelo, L. A., Rademaker, A. W., McConnel, F. M., Heiser, M. A., Cardinale, S., Shedd, D., Stein, D., Beery, Q., Myers, E., Lewin, J., Haxer, M., & Esclamado, R. (1998). Surgical variables affecting speech in treated patients with oral and oropharyngeal cancer. *The Laryngoscope, 108*(6), 908–916. https://doi.org/10.1097/00005537-199806000-00022

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rebernik, T., Jacobi, J., Jonkers, R., Noiray, A., & Wieling, M. (2021). A review of data collection practices using electromagnetic articulography. *Laboratory Phonology, 12*(1), 6. https://doi.org/10.5334/labphon.237

Rebernik, T., Jacobi, J., Tiede, M., & Wieling, M. (2021). Accuracy assessment of two electromagnetic articulographs: Northern Digital Inc. WAVE and Northern Digital Inc. VOX. *Journal of Speech, Language, and Hearing Research, 64*(7), 2637–2667. https://doi.org/10.1044/2021_JSLHR-20-00394

Rinkel, R. N., Verdonck-de Leeuw, I. M., van Reij, E. J., Aaronson, N. K., & Leemans, C. R. (2008). Speech Handicap Index in patients with oral and pharyngeal cancer: Better understanding of patients' complaints. *Head & Neck, 30*(7), 868–874. https://doi.org/10.1002/hed.20795

Roach, P. (2004). British English: Received pronunciation. *Journal of the International Phonetic Association, 34*(2), 239–245. https://doi.org/10.1017/S0025100304001768

Savariaux, C., Badin, P., Samson, A., & Gerber, S. (2017). A comparative study of the precision of Carstens and Northern Digital Instruments electromagnetic articulographs. *Journal of Speech, Language, and Hearing Research, 60*(2), 322–340. https://doi.org/10.1044/2016_JSLHR-S-15-0223

Schuster, M., & Stelzle, F. (2012). Outcome measurements after oral cancer treatment: Speech and speech-related aspects—An overview. *Oral and Maxillofacial Surgery, 16*(3), 291–298. https://doi.org/10.1007/s10006-012-0340-y

Song, Y., Chen, G., & Song, Y. (1984). The free thigh flap: A new free flap concept based on the septocutaneous artery. *British Journal of Plastic Surgery, 37*(2), 149–159. https://doi.org/10.1016/0007-1226(84)90002-X

Speksnijder, C. M., van der Bilt, A., van der Glas, H. W., Koole, R., & Merkx, M. A. W. (2011). Tongue function in patients treated for malignancies in tongue and/or floor of mouth; a one year prospective study. *International Journal of Oral and Maxillofacial Surgery, 40*(12), 1388−1394. https://doi.org/10.1016/j.ijom.2011.09.003

Stone, M., Langguth, J. M., Woo, J., Chen, H., & Prince, J. L. (2014). Tongue motion patterns in post-glossectomy and typical speakers: A principal components analysis. *Journal of Speech, Language, and Hearing Research, 57*(3), 707–717. https://doi.org/10.1044/1092-4388(2013/13-0085)

Takatsu, J., Hanai, N., Suzuki, H., Yoshida, M., Tanaka, Y., Tanaka, S., Hasegawa, Y., & Yamamoto, M. (2017). Phonologic and acoustic analysis of speech following glossectomy and the effect of rehabilitation on speech outcomes. *Journal of Oral and Maxillofacial Surgery, 75*(7), 1530–1541. https://doi.org/10.1016/j.joms.2016.12.004

Teplansky, K. J., Wisler, A., Green, J. R., Heitzman, D., Austin, S., & Wang, J. (2023). Measuring articulatory patterns in amyotrophic lateral sclerosis using a data-driven articulatory consonant distinctiveness space approach. *Journal of Speech, Language, and Hearing Research, 66*(8S), 3076–3088. https://doi.org/10.1044/2022_JSLHR-22-00320

Tienkamp, T. B., van Son, R. J. J. H., & Halpern, B. M. (2023). Objective speech outcomes after surgical treatment for oral cancer: An acoustic analysis of a spontaneous speech corpus containing 32.850 tokens. *Journal of Communication Disorders, 101,* Article 106292. https://doi.org/10.1016/j.jcomdis.2022.106292

Tschiesner, U., Sabariego, C., Linseisen, E., Becker, S., Stier-Jarmer, M., Cieza, A., & Harreus, U. (2013). Priorities of head and neck cancer patients: A patient survey based on the brief ICF core set for HNC. *European Archives of Oto-Rhino-Laryngology, 270*(12), 3133–3142. https://doi.org/10.1007/s00405-013-2446-8

Van den Steen, L., Van Nuffelen, G., Guns, C., De Groote, M., Pinson, L., & De Bodt, M. (2011). De spraak handicap index: Een instrument voor zelfevaluatie bij dysartriepatiënten [The Speech Handicap Index: A tool for self-evaluation for individuals with dysarthria]. *Logopedie, 24,* 26–30.

van Dijk, S., van Alphen, M. J. A., Jacobi, I., Smeele, L. E., van der Heijden, F., & Balm, A. J. M. (2016). A new accurate 3D measurement tool to assess the range of motion of the tongue in oral cancer patients: A standardized model. *Dysphagia, 31*(1), 97–103. https://doi.org/10.1007/s00455-015-9665-7

Wang, J., Green, J. R., Samal, A., & Yunusova, Y. (2013). Articulatory distinctiveness of vowels and consonants: A data-driven approach. *Journal of Speech, Language, and Hearing Research, 56*(5), 1539–1551. https://doi.org/10.1044/1092-4388(2013/12-0030)

Weismer, G. (2013). Speech science: Technique, concept, theory. *Perspectives on Speech Science and Orofacial Disorders, 23*(1), 6–17. https://doi.org/10.1044/ssod23.1.6

Weismer, G., Yunusova, Y., & Bunton, K. (2012). Measures to evaluate the effects of DBS on speech production. *Journal of Neurolinguistics, 25*(2), 74–94. https://doi.org/10.1016/j.jneuroling.2011.08.006

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Yang, G.-F., Chen, P.-J., Gao, Y.-Z., Liu, X.-Y., Li, J., Jiang, S.-X., & He, S.-P. (1997). Classic reprint forearm free skin flap transplantation: A report of 56 cases. *British Journal of Plastic Surgery, 50*(3), 162–165. https://doi.org/10.1016/S0007-1226(97)91363-1

Zhou, X., Woo, J., Stone, M., & Espy-Wilson, C. (2013). A cine MRI-based study of sibilant fricatives production in post-glossectomy speakers. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 7780–7784). https://doi.org/10.1109/ICASSP.2013.6639178

Appendix

The North Wind and the Sun Passage

Dutch text as included in the present study
De noordenwind en de zon waren erover aan het redetwisten wie de sterkste was van hen beiden. Juist op dat moment kwam er een reiziger aan, die gehuld was in een warme mantel. Ze waren het erover eens dat degene die er als eerste in slaagde de reiziger zijn mantel uit te doen, als sterker moest worden beschouwd dan de ander. De noordenwind begon toen uit alle macht te blazen. Maar hoe harder hij blies, des te dichter trok de reiziger zijn mantel om zich heen. Ten lange leste gaf de noordenwind het op. Daarna begon de zon krachtig te stralen, en hierop trok de reiziger onmiddellijk zijn mantel uit. De noordenwind moest dus wel bekennen dat de zon van hen beiden de sterkste was.
English version of Roach (2004)
The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.