Consonant production in Greek Lombard speech: an electropalatographic study

Katerina Nicolaidis

Speech production in noise is characterised by important articulatory and acoustic modifications. The current study examines lingual articulation during consonant production in speech produced in quiet and noise. It uses the technique of electropalatography to investigate lingual-palatal contact patterns during the production of the Greek consonants /t, k, s, x, n, l, ɾ/ in the two conditions. Two male and two female speakers produced words embedded in carrier phases in quiet and in multitalker babble noise played over loudspeakers. Several parameters are analysed including consonantal duration, total amount of contact in the alveolar and palatal regions, place of articulation for all consonants, degree of constriction for the fricatives, V-to-C coarticulatory effects and token-to-token variability. The results provide evidence of spatio-temporal modifications in speech produced in noise. These include an overall tendency for shorter duration for all consonants, more lingual-palatal contact and more advanced placement in the alveolar zone as well as reduced contact in the palatal region for /t, s, n, l, ɾ/, and smaller coarticulatory effects. Variability to these patterns was however evident. Variability as a function of speaker, gender, consonant, vowel and condition is examined. The results are discussed within the framework of adaptive variability proposed by the Hyper-Hypo Speech model (Lindblom 1990). They suggest that in addition to changes potentially related to increased vocal effort there is some first evidence pointing towards speech modifications that can be communicatively driven.*

1. Introduction

1.1. Theoretical background

Speech communication among individuals frequently occurs in non optimal conditions, such as the presence of noise. Noise can have a detrimental effect to communication since it can mask important acoustic features of speech. Research has shown that speech production in noise adapts along several parameters to compensate for this masking effect. Modifications in articulatory and acoustic parameters are known as the Lombard effect (Lombard 1911) and

Rivista di Linguistica 24.1 (2012), pp. 65-101 *(received November 2011)*

include an increase in amplitude, fundamental frequency, word and vowel duration, a shift in F1 and F2 frequencies (mainly the former), flatter spectra with a shift of energy from low to middle or high frequency bands as well as an increase in the amplitude of movements (e.g., Dreher & O'Neill 1957, Pisoni et al. 1985, Van Summers et al. 1988, Bond et al. 1989, Junqua 1993, 1996, Letowski et al. 1993, Castellanos et al. 1996, Steeneken & Hansen 1999, Pittman & Wiley 2001, Kim 2005, Varadarajan & Hansen 2006, Garnier et al. 2006, Davis et al. 2006).

The mechanisms underlying the modifications that occur in Lombard speech, i.e., purely physiological, cognitive or both, have been a challenging issue to address. Different accounts have been put forward proposing that changes are due to (a) a physiological audio-phonatory reflex leading to an automatic change of vocal intensity due to modifications in auditory feedback (Lombard 1911); (b) an active reorganisation of speech aiming towards maintaining or improving the intelligibility of speech for communicative functionality (Lane & Tranel 1971); (c) a combination of both, that is, an automatic response and a communicative adaptation (Junqua 1993, Garnier et al. 2006, Garnier et al. 2010).

The study of Lombard speech is thus important both from a theoretical standpoint for the development of comprehensive model that predicts and accounts for the modifications that occur, as well as for practical applications such as the development of robust speech recognition systems for use in different and demanding contexts, e.g., aircraft control, the development of cochlear implants and conventional hearing aids.

Whatever the mechanism(s) underlying the speech modifications that occur in Lombard speech, it is evident that these changes are the outcome of the link between perception and production. The increase of vocal intensity can enhance auditory feedback for the speaker but can also make the speech signal more audible for the listener. Moreover, perceptual studies have shown that speech produced in noise is more intelligible than speech produced in quiet conditions (Dreher & O'Neill 1957, Pittman & Wiley 2001, Van Summers et al. 1988, Lu & Cooke 2008). Further evidence showing that vocal intensity increases more in conversational speech than reading in noise suggests that communicative involvement plays an important role in the modifications that take place (Amazi & Garber 1982, Junqua et al. 1999). Garnier et al. (2010) report greater increase of several acoustic parameters and further speech modifications present only in interactive conditions in noise. Taken together, such studies indicate that articulatory adaptations can function to enhance the intelligibility of speech in noise and can aim towards communicative functionality. Depending on the degree of communicative involvement of the interlocutors and the environmental noise conditions, variability may be expected to occur in the articulatory movements of the speaker and the resulting acoustic signal. Indeed, one of the features that consistently stands out in previous research on speech production in noise is the large variability present. Junqua (1996: 16) describes Lombard speech as a phenomenon "distributed along a continuum of speech variability which is influenced by a complex set of parameters such as the nature of the speakers, the context and the environment".

Adaptive variability guided by the demands of the communicative situation is a key principle of the Hyper-Hypo-speech (H&H) model (Lindblom 1990). According to the model, speech production is adaptive and ranges between a continuum of hyper-hypo speech. Two fundamental principles are inherent in the model: (a) system-oriented constraints which are characterised by motor control optimisation and economy, and result in hypo-forms; (b) output-oriented constraints which reflect the teleological nature of speech production and necessitate adjustments for communicative functionality. These produce variability effects that are listener-oriented and are characterised by hyper-forms. It is argued that there is a universal tendency for the predominance of system-oriented constraints, i.e., for speakers to hypo- rather than hyper-articulate. However, this is implemented to the degree that is communicatively purposeful thus rendering output-oriented constraints fundamental in shaping intra-speaker variation. The aim of the speaker is to meet the demands of the communicative environment and vary his/her articulation so as to increase efficiency in communication. During speech produced in noise, it can be assumed that output-oriented constraints are severe and speakers modify their articulatory strategies in order to maximise discriminability and intelligibility for effective communication. An active reorganisation of articulatory movements can thus occur resulting in the production of hyper-forms. It can be predicted that modifications may result in increases in intensity, vowel duration, frequency changes, amplitude of movements and so on.

1.2. Factors influencing variability in Lombard Speech

Previous studies on Lombard speech have reported variability in several acoustic and articulatory parameters as a function of various factors including the speaker, the noise type and level, the language and the situational context (see also section 1.1). One of the main factors that has been shown to contribute to the high variability of Lombard speech is the speaker (e.g., Junqua 1993, Garnier et al. 2010) as different articulatory strategies may be employed by speakers when increasing their vocal effort. In addition, variability may relate to the speaker's intentions and willingness to communicate in a particular situation. Noise also exerts a psychological effect on the speaker, and different responses, inducing variable acoustic changes, may be expected by different individuals. Important gender-related differences have also been observed in terms of variability in pitch level and formant frequencies between male and female speakers (Anglade & Junqua 1990, Junqua 1993, Letowski et al. 1993, Castellanos et al. 1996).

The Lombard effect is also dependent on the type and level of noise. Junqua (1994) has shown, for example, that vowel duration increases more in the presence of multi-talker babble noise compared to white-Gaussian noise while Garnier et al. (2006) report increased vowel duration in white noise compared to cocktail party noise. Lu & Cooke (2008) used babble noise varying in number of speakers and found that an increase in noise level and the number of background talkers led to an increase in several acoustic parameters. Sound immersion techniques have also been found to influence the extent of the modifications present in noise. Garnier et al. (2010) report greater changes in several acoustic parameters when cocktail party noise was played over headphones than loudspeakers. Davis et al. (2006) found greater amplitude of articulatory movements in the headphone condition compared to loudspeakers.

Language-specific variability has also been studied in languages such as English, French, Spanish, Japanese, Korean, Greek (e.g., Junqua 1993, Lu & Cooke 2008, Ramez 1992, Garnier et al. 2010, Castellanos et al. 1996, Takizawa & Hamada 1990, Kim 2005, Nicolaidis & Rispoli 2005). Junqua (1996) summarises some of the findings of previous studies on American English, French, Spanish, and Japanese and discusses similarities and differences in the tendencies observed. Other linguistic variables are also important, such as type of word (content vs. function words) and syllabic position. For example, larger effects on selected acoustic and articulatory parameters have been reported for content than function words in noise (Patel & Schell 2008, Garnier et al. 2006; the latter study also examines syllable position effects). Enhancement of prosodic cues in noise has also been reported (Garnier et al. 2006, Welby 2006, Garnier et al. 2010).

To date, the majority of studies have focused on the examination of acoustic parameters. Relatively limited articulatory work is available and little is known about the articulatory modifications that occur in Lombard speech, how the different articulatory subsystems function and interact, how articulatory changes relate to acoustic changes and how the variables expounded above affect articulatory variability. The need for more physiological studies for speech in adverse conditions is also underscored in Van Summers et al. (1989). Furthermore, the majority of studies on segmental aspects have mainly focused on vowels leaving consonant production largely unexplored.

The articulatory studies currently available have mainly focused on selected labial parameters as well as on visual enhancement in noise. They have reported amplification of articulatory movements (lip aperture and spreading), changes in velocity peaks relating to articulatory effort (Garnier et al. 2006), increased lip compression for bilabial stops (Garnier 2008, Garnier et al. 2010) and articulatory and durational changes on selected units in utterances (Patel & Schell 2008, Garnier et al. 2006) functioning to enhance cues to word segmentation and prosodic hierarchy, and thus intelligibility, in noise.

1.3. Aim of the current study

This study investigates the effect of noise on the lingual system during consonant production using the technique of electropalatography. Several parameters are examined including the total amount of linguo-palatal contact in different regions of the palate, place of articulation changes, V-to-C coarticulatory effects, and token-to-token variability during the production of the Greek consonants /t, s, k, x, n, l, ɾ/ in quiet and multi-talker babble noise. Variability as a function of speaker, gender, consonant, vowel and condition (quit noise) is examined. The study aims to provide articulatory data on the lingual system and add to the current knowledge-base pertaining to variability in Lombard speech on the basis of mainly acoustic research on other languages. Some of the major questions it addresses include:

Is there a difference in amount of lingual contact between the quiet and noise conditions?

Are there differences in the place of articulation of consonants in the two conditions?

How can such potential differences be interpreted (i.e., result of increase in vocal effort vs. speaker controlled adaptive variability)?

Are there differences in the degree and nature of coarticulatory effects present?

Are there differences in the amount of token-to-token variability suggesting differential degree of articulatory precision in the two conditions?

From both a theoretical and methodological perspective, this study examines many of the key research topics addressed in this thematic issue. It analyses articulatory (EPG) data to investigate variability in phonetic form and coarticulation examined within the framework of the H&H theory. It also examines speaker-specific and gender-related variation. The study of Lombard speech can add to the expanding agenda of socio-phonetic research by examining variation in a particular case of contextual interaction, i.e., during communication in adverse conditions. Variation is accounted for by investigating core parameters such as the link between speech perception and production, articulatory constraints, and speaker-specific strategies and preferences determined to a large extent by the varying demands of the situational context. An interplay between bio-physical, psychological, linguistic, environmental and social factors is present in Lombard speech; all these shape variability.

2. Methodology

2.1. Subjects

Two male (AT, TP) and two female (MM, EP) Greek speakers were recorded for this study. They were between 38-46 years of age and lived in Thessaloniki. They were speakers of Standard Modern Greek, i.e., with no regional accents. All of them had normal hearing and no speech defects. They were all naïve with regard to the purposes of the experiment.

2.2. Speech Material

The speech material consisted of disyllabic words of the form C_1VC_2V. The intervocalic consonants that were analysed were /t, k, s, x, n, l, ɾ/, the initial consonant was one of the labial consonants /p, b, f/, and V = /i, a/. A description of the Greek consonantal and vocalic system can be found in Arvaniti (1999). Symmetrical sequences were recorded with stress on the first syllable, e.g., /ˈpapa/, /ˈbala/, /ˈfaka/, /ˈpasa/. An initial labial consonant was selected as labial closure involves an independent system and thus influence on the lingual gestures is minimised. In addition, the initial labial consonant varied among /p, b, f/ so as to enable the inclusion of real words in the corpus. The speech material also included one proper name (i.e., /ˈmaɾa/) and just one nonsense item (i.e., /ˈpaxa/) which was presented to the subjects with a capitalised first letter to resemble a proper name. Test words were embedded in the carrier phrase /ˈleɣe ˈCVCV ˈpali/

('say___again'). Subjects were instructed to read the speech material at a comfortable rate. They initially produced all material in a quiet condition and then in a multi-talker babble noise condition. For the noise condition, subjects were instructed to read the material as if trying to communicate the information to the experimenter who was present in the room. Five repetitions of all speech material in each condition were analysed. A total of 560 tokens were analysed for this study (7 consonants x 2 vowels x 5 repetitions x 4 speakers x 2 conditions (quiet/noise)).

It should be noted that this investigation is part of ongoing research investigating all Greek consonants in different vocalic environments in words produced in carrier phrases and in spontaneous speech in quiet and noise conditions. The present study includes a selection of the Greek consonants in two different vocalic environments produced in carrier phrases. Clearly, even taking into consideration the instructions given to the subjects, reading a list of words does not involve communicative interaction and thus the full spectrum of speech adaptations may not occur (Garnier et al. 2010). Still modifications may be expected to be present and the analysis of the current, more controlled material will provide a reference for future work on spontaneous speech in noise that will enable the identification of potential differences between the two tasks and provide insights on the role of communicative involvement in Lombard speech.

2.3. Data acquisition

Electropalatographic and acoustic data were simultaneously recorded using the British EPG system marketed by Articulate Instruments. The artificial palate used in this system has 62 electrodes on its surface that record lingual contact with the palate in continuous speech. The electrodes are distributed in eight rows and correspond to particular articulatory regions, i.e., the front four correspond to the alveolar zone and the back four to the palatal zone (Figure 1). The alveolar zone is subdivided to the alveolar and postalveolar regions (rows 1 to 2 and 3 to 4 respectively) while the palatal zone is subdivided to the pre-, medio- and post-palatal regions. The most posterior row is located on the junction between the hard and soft palates so some limited contact for velar sounds is expected to be registered on this row. The first two columns on the left and right side of the palate are characterized as lateral and the four remaining columns as central (Recasens et al. 1993; Gibbon & Nicolaidis 1999). EPG data were sampled at 10 ms.

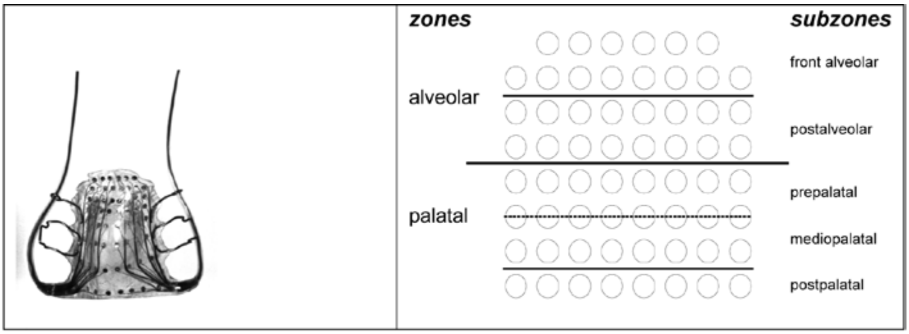


Figure 1. The artificial palate used in this study (left). Schematic representation of electrodes and division of electrodes into zones and subzones (right).

Data were recorded in a quiet and in a multi-talker babble noise condition. For the quiet condition, data were recorded in a sounded-treated room in the Phonetics laboratory of the School of English at Aristotle University. A Sennheiser e815S microphone was used. For the noise condition, data were recorded in a sound treated recording studio at the laboratory of Electroacoustics & TV Systems of the Department of Electrical and Computer Engineering at Aristotle University. The babble noise was created from 8 mixed voices (four male and four female) and was further processed to remove stretches that were comprehensible so that the resulting babble noise was completely unintelligible. Noise was calibrated at 88dB SPL at the speaker's ears played over two loudspeakers spaced three meters apart and at a distance of three meters from the subject. Sound immersion with noise played over loudspeakers rather than headphones was opted for in order to simulate a more naturally occurring condition for the speaker. The audio signal was recorded with a cardiod Beyerdynamic MC836PV microphone placed at a distance of 5 cm from the speaker' mouth. Important acoustic landmarks, e.g., onset and end of formant structure, evidence of bursts, could be identified in the resulting acoustic signal so it was used for segmentation and annotation without noise extraction and in conjunction with information obtained from the electropalatographic data.

2.4. Data measurement and analysis

The software provided with the EPG system, Articulate Assistant (version 1.18), was used for data display, segmentation and analysis. For the duration measurements reported in the study, segmentation was based on the acoustic waveform and spectrogram. Onset of consonantal constriction was taken at the end of the formant structure for the preceding vowel. End of consonantal constriction was taken at the onset of formant structure for the following vowel. For the stops /t, k/, end of consonantal constriction was taken at the burst. For some /ɾ/ tokens there was evidence of release on the acoustic data and end of constriction was taken at that point. Segmentation was facilitated from the EPG data when acoustic information was limited in the noise condition. For the fricatives /s, x/, onset and offset was annotated at the onset/offset of noise.

For the articulatory analysis, the first frame of maximum contact on the entire palate, i.e., all eight rows, was annotated for all consonants (Figure 2). For the consonants /t, s, n, l, ɾ/ this typically coincided with maximum contact in the alveolar zone (front four rows) and for the /k, x/ in the palatal zone (back four rows). In the few cases that it did not, the first frame of maximum contact in the alveolar or palatal region was selected during the interval of maximum contact on the entire palate for the consonant. For the fricatives, the first frame of maximum constriction in the alveolar zone for /s/ and the palatal zone for /x/ during the interval of maximum contact on the entire palate was annotated. As the first frame of maximum contact/constriction can vary temporally, the EPG frame at the temporal midpoint of the consonantal interval was also extracted from the data for the examination of coarticulatory effects at the consonantal midpoint.

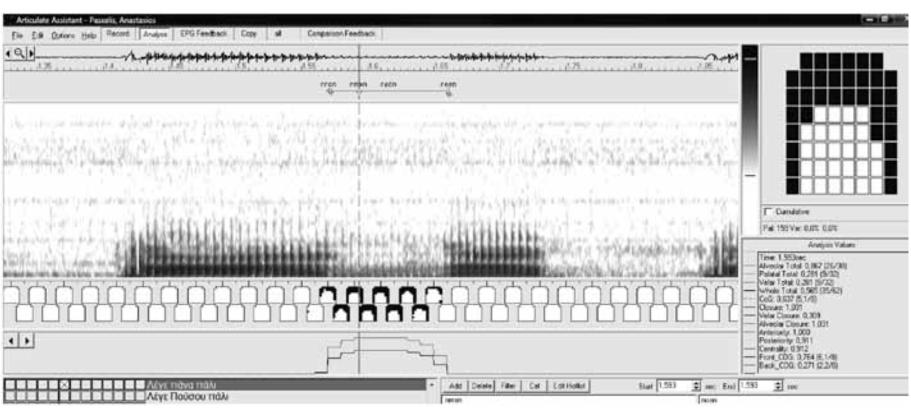


Figure 2. Screen display and segmentation of the consonant /n/ in the word /'pana/ produced in noise. The segmentation line corresponds to the first frame of maximum contact which is displayed at the top right of the figure. The display below the spectrogram and palatograms shows the total amount of contact at the alveolar zone (top line) and the entire palate (bottom line).

EPG data reduction methodology was based on the calculation of the following measures at the frame of maximum contact/constriction:

(a) the percentage frequency of electrode activation of the entire palate over five repetitions,

(b) the total number of contacts for the anterior four rows (front total) for the consonants /t, s, n, l, ɾ/,

(c) the total number of contacts for the posterior four rows (back total) for all consonants,

(d) the total number of contacts for the entire palate, i.e., all eight rows, (global total) for the palatal productions of /k, x/ in the environment of /i/,

(e) the Centre of Gravity for the front four rows (front CoG) for the consonants /t, s, n, l, ɾ/,

(f) the Centre of Gravity for the back four rows (back CoG) for the consonants /k, x/,

(g) the Centre of Gravity for the entire palate (global CoG) for the palatal productions of /k, x/ in the environment of /i/,

(h) the mean lateral measure for the front four rows (front lateral) for /s/,

(i) the mean lateral measure for the back four rows (back lateral) for /x/,

(j) a variability index for the entire palate.

The totals measures and contact indices are described in Gibbon & Nicolaidis (1999) and the Articulate Assistant User Guide. The totals measures were ratios of the number of contacted electrodes over the total number of electrodes in the region.

The Centre of Gravity index expresses the location of the main concentration of activated electrodes. For the calculation of the index, progressively higher values are assigned to more anterior rows. Thus a higher value of the index indicates more anterior placement for a particular sound. The mean lateral measure was used to show if there was more contact near the midline or the lateral sides of the palate. The index provides a measure of the constriction degree for the fricative sounds. A higher value of the index indicates more central articulation, i.e., greater constriction. For the consonants /t, s, n, l, ɾ/, relevant indices were calculated for the front four rows because their place of articulation always occurred in the alveolar zone. For /k, x/, the indices were calculated for the back four rows as their place of articulation occurred in the palatal zone. Finally, the variability index quantifies the degree of variability in EPG patterns across repetitions of a particular sound. For a given electrode, 100% and 0% activation across repetitions represent invariance while contact around 50% represents maximum variability (Farnetani & Provaglio 1991). Calculation of the index for different consonants can provide a measure of articulatory constraint and precision during their production.

As a measure of V-to-C coarticulation, the difference of various measures in the /i/ vs. /a/ contexts was calculated both at the frame of maximum contact/constriction and at the temporal midpoint of the consonants. In particular, differences in the total amount of alveolar contact (front total), palatal contact (back total) and the values of the contact indices (front CoG, back CoG, front and back lateral) were calculated.

Factorial analyses of variance on the different measures, i.e., consonant duration, front, back and global total, front, back and global GoG,[1] front and back lateral, and variability index at the frame of maximum contact/constriction and at the temporal midpoint were carried out including subject (MM, EP, AT, TP), consonant (selected consonants from /t, s, k, x, n, l, ɾ/ depending on the measure), vowel context (/i, a/), and condition (quiet/noise) as factors. Separate analyses with gender, consonant, vowel context and condition were also carried out.

3. Results

Adaptations in noise in terms of duration, total amount of linguopalatal contact, location and degree of constriction will be discussed below. Separate sections on the effect of condition, consonant type, vowel context, speaker and gender are included so that important information relevant to general phonetic and sociophonetic studies is presented. The results of the statistical analyses of the duration measurements, and the different articulatory measures calculated at the frame of maximum contact/constriction will be reported. In addition, V-to-C coarticulatory effects at the temporal midpoint will be presented, i.e., the vowel main effect and interactions including vowel and consonant. The reader should note that only these effects will be reported for the temporal midpoint.

EPG palatograms displaying percentage frequency of activation over five repetitions at the frame of maximum contact/constriction can be seen in Figure 3 for all consonants produced in the quiet and noise conditions for one male and one female speaker. The reader can refer to the palatograms for an illustration of the articulatory modifications described in the sections below.

3.1. Effect of condition on consonant production

The ANOVA analysis showed a significant main effect of condition, i.e., quiet vs. noise, on consonant duration ($F(1, 559) = 65.93$, $p < 0.0001$). The means for the consonant by condition interaction ($F(6, 559) = 4.062$, $p < 0.001$) showed that overall consonantal duration was shorter in the noise condition for all consonants; Tukey post-hoc tests showed that significant differences were only found for the consonants /x, n/ (Table 1). The means for the subject by consonant by condition interaction ($F(18, 559) = 3.149$, $p < 0.0001$) showed that overall shorter consonantal duration for consonants in noise was true for all subjects; few exceptions were evident. Figure 4 shows average duration per consonant in the quiet and noise conditions and the percentage difference in duration between the two conditions for all consonants. For the obstruents, greatest compression under noise is evident for the fricative /x/ and the smallest difference between conditions is noted for the stop /t/. For the sonorants, larger differences between conditions were found for /n/ than /l/ while the large percentage difference noted for the tap was mainly due to one speaker (TP).

With reference to the articulatory measures for /t, s, n, l, ɾ/, the analysis of the front total indicated more contact in the alveolar region in noise ($F(1, 399) = 4.425$, $p < 0.036$) (Figure 3); examination of the means of the consonant by condition interaction showed that this was true mainly for the consonants /t, s, ɾ/ ($F(4, 399) = 6.874$, $p < 0.0001$) while post hoc tests showed that significance was only reached for /s/. The consonant by subject by condition interaction revealed speaker variability ($F(11, 399) = 4.473$, $p < 0.0001$). The means showed that the tap was produced with more alveolar contact in noise by all speakers, /t, s, n/ by three speakers and /l/ by two speakers. Subject TP showed less contact for /t, n, l/ in noise with a relatively large difference between conditions for /n/ and especially /l/. It should be noted that speaker TP generally produced very palatalised consonants in the environment of /i/ in the quiet condition. His production of /n/ in /ini/ was categorized as palatal involving increased alveolar and palatal contact. Such palatalisation was reduced in noise, hence the smaller amount of contact evident in the alveolar region (similarly for /l/). In addition, a very palatal production of the /s/ in the environment of /i/ was found for this speaker in quiet while in noise an alveolar pattern was evident leading to a large increase in alveolar contact in this condition. Hence post hoc tests showed significant differences for /s/ (with more contact in noise), as well as /n, l/ (with less contact in noise) only for this speaker.

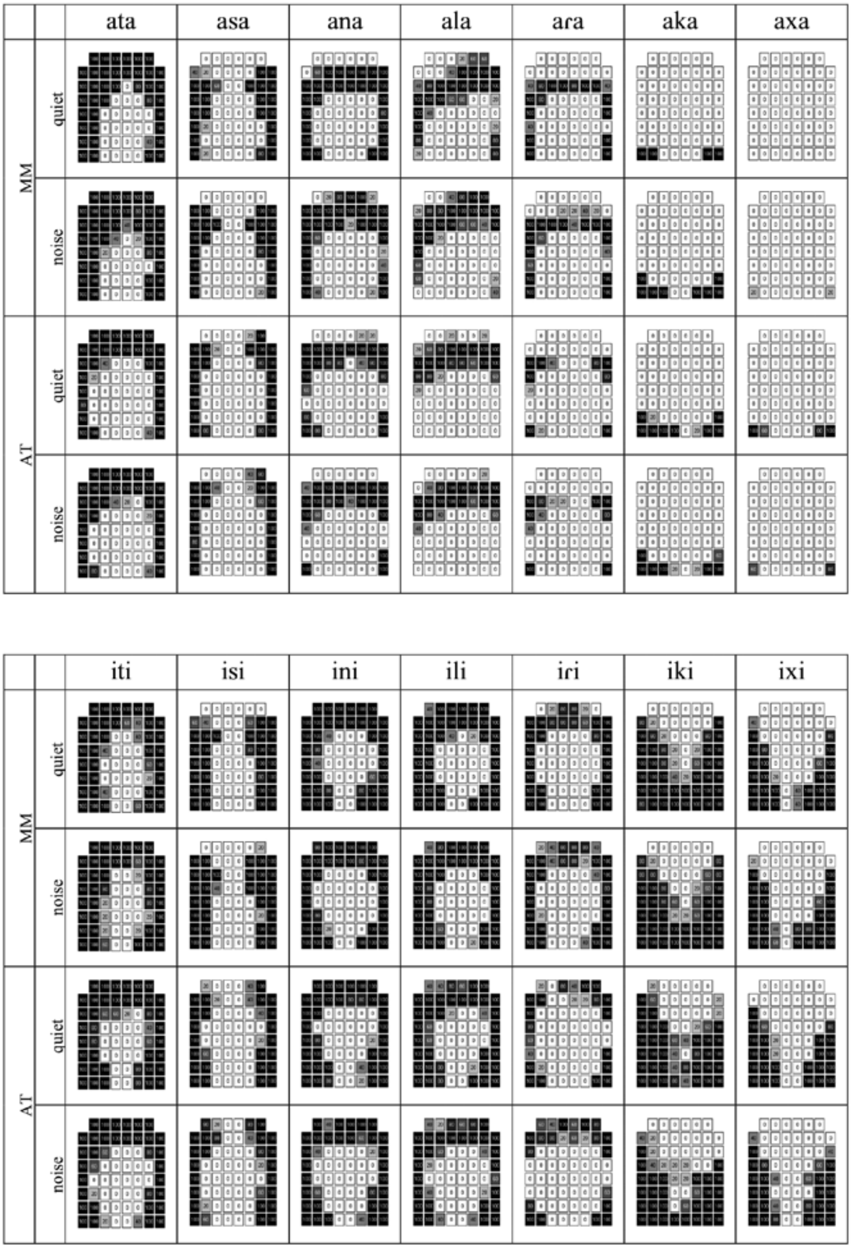
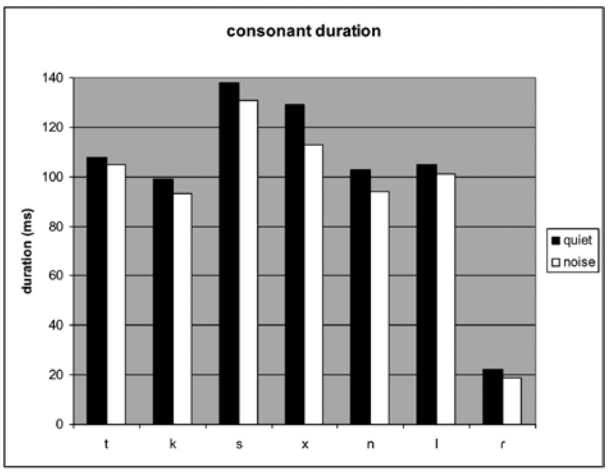


Figure 3. Percentage frequency of activation over 5 repetitions at the frame of max contact/constriction for /t, s, n, l, ɾ, k, x/ in the /a - a/ (top) and /i - i/ (bottom) contexts in the quiet and noise conditions. Palatograms for one female (MM) and one male speaker (AT) are presented.

Table 1. Average durations (in ms) and standard deviations for the consonants /t, k, s, x, n, l, ɾ/ in quiet and noise.

	QUIET		NOISE	
	AVERAGE	ST. DEV	AVERAGE	ST. DEV
t	108	10	105	13
k	99	19	93	15
s	138	19	131	19
x	129	19	113	16
n	103	14	94	15
l	105	16	101	13
ɾ	22	8	19	4



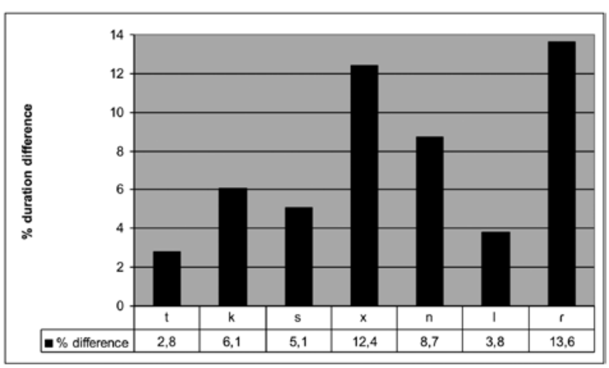


Figure 4. Average duration (in ms) for the consonants /t, k, s, x, n, l, ɾ/ in the quiet and noise conditions (upper graph). Percentage difference in duration between the quiet and noise conditions for the consonants examined (bottom).

The analysis of the back total showed overall less contact in the noise condition (F(1, 399) = 165.954, p < 0.0001) (Figure 3). Significant differences for all consonants (consonant by condition: F(4, 399) = 3.569, p < 0.007) were found in the post hoc tests. Examination of the means of the consonant by subject by condition interaction (F(12, 399) = 3.831, p < 0.0001) showed that this pattern was observed in all cases with only very few exceptions present.[2] Post-hoc tests showed that differences reached significance for all consonants for speakers TP and AT (but for /ɾ/), and MM for /n/.

With reference to place of articulation, the analysis of the front CoG showed overall more anterior placement in the noise condition (F(1, 399) = 43.443, p < 0.0001); the means showed that this occurred for all consonants but for /t/ which was similarly produced in the two conditions (F(4, 399) = 5.818, p < 0.0001) (Figure 3). Post-hoc tests showed significant differences for /n, l, ɾ/ (Figure 5).

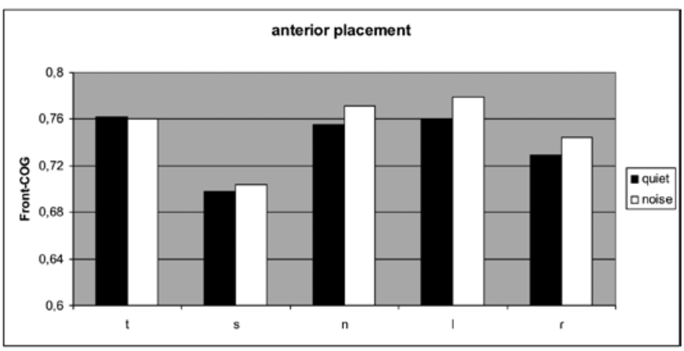


Figure 5. Front Centre of Gravity for the consonants /t, s, n, l, ɾ/ in the quiet and noise conditions. The higher the index the more anterior the place of articulation of the consonant.

The means for the subject by consonant by condition interaction showed that this was true in the majority of cases but for a few exceptions (F(12, 399) = 7.386, p < 0.0001).[3] Post-hoc tests showed that significant differences were evident for /n, l, ɾ/ for three speakers and for two speakers for /s/. For speaker AT no significant differences were found for any consonant.

In addition, for the fricative /s/, the results of the front lateral measure showed significantly more central contact in the noise condition (F(1, 79) = 23. 693, p < 0.0001) (Figure 3).

For the consonants involving primary constriction in the palatal/velar zone, i.e., /k, x/, the analysis of the back total showed that contact did not vary significantly between the quiet and noise conditions. Similarly, for the back CoG measure, no significant differences in lingual placement in the palatal zone in the two conditions were found (see Figure 3). As contact is evident in more anterior regions for the palatal production of /k, x/ in the environment of /i/ (Figure 3), separate analyses on the global total and global CoG were carried out. The main effect of condition did not reach significance indicating that production did not differ in the two conditions. Finally, condition did not induce any significant effects on the degree of constriction of the consonant /x/ as shown by the results of the back lateral measure (Figure 3). It should be noted that although some contact is registered for the velars /k, x/ in the posterior rows of the artificial palate, potential differences in placement or degree of constriction for /x/ occurring further behind this area cannot be revealed by the EPG data.

3.2. Effect of consonant type

Duration varied significantly as a function of consonant ($F(6, 559) = 1089.55$, $p < 0.0001$) with the fricatives being longest and the rhotic the shortest (Figure 4a). Post hoc tests showed that all duration differences among consonants were significant except for /t/ and /l/, and /k/ and /n/.

The analysis of the front and back total for the consonants /t, s, n, l, ɾ/ showed significant differences in amount of contact among the consonants in both articulatory regions (front total: $F(4, 399) = 10055.659$, $p < 0.0001$; back total: $F(4, 399) = 549.904$, $p < 0.0001$). The total amount of contact in the alveolar zone decreased in the order t>n>l>s>ɾ and in the palatal zone in the order t>s>n>l>ɾ; post-hoc tests showed that all differences among consonants were significant. In addition, the analysis of the front CoG showed significant differences in anterior placement among these consonants ($F(4, 399) = 251.174$, $p < 0.0001$). The consonant by subject interaction ($F(12, 399) = 72.080$, $p < 0.0001$) showed variability in placement among speakers with a general trend for more anterior placement for /l, n, t/, and less so for /ɾ/ and /s/.

Finally, for the back consonants /k, x/, the consonant main effect was significant for the back total measure ($F(1, 159) = 661.376$, $p < 0.0001$) and the back CoG ($F(1, 159) = 253.100$, $p < 0.0001$). More contact in the palatal zone was present for the stop than the fricative; this difference was significant for all speakers but EP ($F(3, 159) = 42.162$, $p < 0.0001$). In addition, more fronted production was found for the stop compared to the fricative for all speakers ($F(3, 159) = 11.948$, $p < 0.0001$) (Figure 3).

3.3. Effect of vowel context

A significant vowel effect on consonantal duration was found with longer duration present in the /i/ than /a/ context (vowel: (F(1, 559) = 161.88, p < 0.0001, consonant by vowel by condition: (F(6, 559) = 2.42, p < 0.026); post hoc tests showed that for the fricatives this difference was significant in both the noise and quiet conditions while for /k, n, l/ significant differences were only evident in the quiet condition. For /t, ɾ/ effects did not reach significance in either condition. Similar inherent duration (section 3.2) and context-induced differences for selected consonants and environments have been reported before for Greek speech produced in quiet (e.g., Nicolaidis 2001, 2002, see Arvaniti 2007 for a review).

For the front total measure analysed for /t, s, n, l, ɾ/, the vowel main effect showed significant influence of the context with overall more contact during consonant production in the environment of /i/ (F(1, 399) = 55.045, p < 0.0001). However, the consonant by subject by vowel interaction (F(12, 399) = 5.128, p < 0.0001) showed variability in the vowel effects, e.g., more contact in the alveolar region was generally evident in the context of /a/ for /t/. Post-hoc tests showed that not all differences were significant. Examination of vowel effects at the temporal midpoint showed similar tendencies (subject by consonant by vowel: (F(12, 399) = 5.190, p < 0.0001). The means of the vowel by condition interaction showed a tendency for smaller V-to-C effects in noise, i.e., smaller difference in amount of alveolar contact between the /i/ and /a/ context in the noise than quiet condition. Such a difference was not however statistically significant.

Analysis of the back total measure for /t, s, n, l, ɾ/ showed more contact in the /i/ context (F(1, 399) = 4.114, p < 0.0001) in both conditions F(1, 399) = 78.363, p < 0.0001). The means for the consonant by vowel by condition interaction (F(4, 399) = 4.608, p < 0.001) showed consistently smaller coarticulatory effects in noise compared to quiet for all consonants (Figure 6 upper graph). This was evident in the difference in palatal contact in the /i/ vs. /a / contexts in the two conditions. In addition, coarticulatory effects decreased in the order l>n>t>r>s in both conditions although in the noise condition small and similar effects were evident on /t/, /s/ and /ɾ/.

Similarly, at the temporal midpoint V-to-C effects were smaller in the noise condition for all consonants but for /s/ which showed similar effects in the two conditions. Effects decreased in the order l>n>t>r>s in quiet and l>n>r>t>s in noise; again in the noise condition, small and similar effects were evident on /t/, /s/ and /ɾ/ (F(4, 399) = 6.907, p < 0.0001) (Figure 6 bottom).

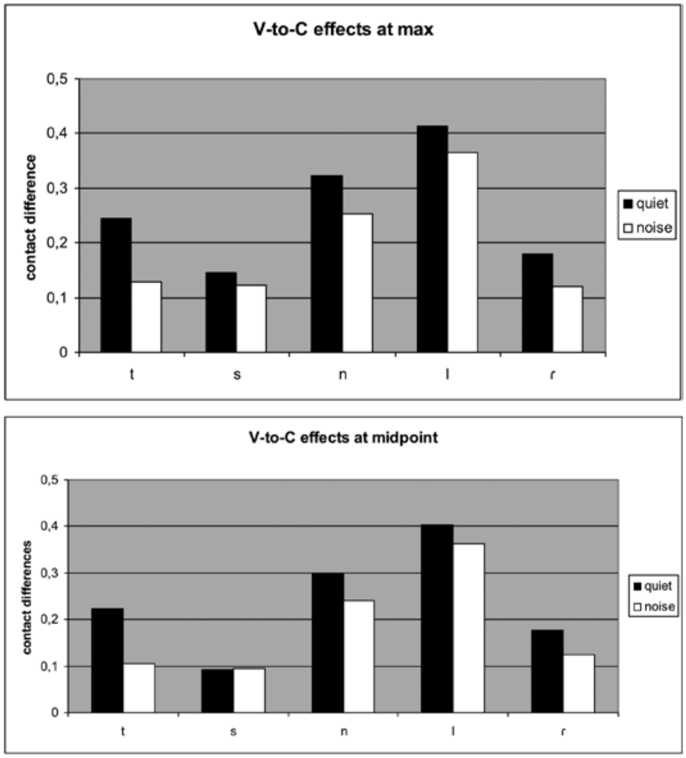


Figure 6. V-to-C effects on the tongue dorsum for the consonants /t, s, n, l, ɾ/ in the quiet and noise conditions at the frame of maximum contact/constriction (upper graph) and the temporal midpoint frame (bottom). The graphs plot the differences in the back total measure, i.e., the total amount of contact in the palatal region (four posterior rows of the EPG palate) during the consonant in the environment /i/ vs. /a/ in the two conditions.

Such coarticulatory findings are in agreement with the model of lingual coarticulation based on articulatory constraints (DAC model) proposed by Recasens et al. (1997). The model predicts coarticulatory effects on the basis of the degree of tongue dorsum involvement in closure or constriction formation, interarticulatory coupling and manner of articulation requirements. Similarly to findings from other languages (see Recasens & Espinosa 2009 for a review), our data show smallest effects on /s/ due to large articulatory constraints involved in the production of this sound for the formation of constriction as well as aerodynamic constraints. The largest effects were evident on /l/ and /n/; the tongue dorsum is not involved in the formation of constriction for either sound while the reduced amount of linguo-palatal contact in the palatal region results in smaller degree of articulatory constraint.

Furthermore, the analysis of the front CoG for /t, s, n, l, ɾ/ showed overall more fronted contact in the environment of /i/ ($F(1, 399) = 554.710$, $p < 0.0001$). This was true of all consonants ($F(4, 399) = 169.649$, $p < 0.0001$); post-hoc tests showed significant differences for /n, l, ɾ/. Some speaker variability was evident (consonant by subject by vowel condition: ($F(15, 399) = 5.498$, $p < 0.0001$)) as in, for instance, more fronted contact in the /a/ context for /t/ for speakers AT and EP. Coarticulatory effects were generally large for the liquids and nasal (decreasing in the order r>n>l) and small on /s/ and /t/. The same was evident at the temporal midpoint ($F(4, 399) = 114.599$, $p < 0.0001$). The means for the consonant by vowel by condition interaction showed that smaller V-to-C effects were overall evident in the noise condition for /n, l, ɾ/, i.e., there were smaller differences in the two vowel contexts. For /t, s/, V-to-C effects were small in both conditions and relatively larger in noise ($F(4, 399) = 5.855$, $p < 0.0001$). Similar effects were evident at temporal midpoint (Figure 7).

Strong articulatory constraints on the tongue tip/blade articulator for the formation of constriction during /s/ (formation of central groove) and /t/ (lingual bracing requiring great degree of linguo-palatal contact) may account for the small contextual effects present (cf. Nicolaidis 1999). It should be noted that the greatest amount of contact in both the alveolar and palatal regions was found for /t/. The tap was found to be produced with the smallest amount of contact in the alveolar and palatal regions and to allow greatest contextual effects along the horizontal anterior/posterior axis in the alveolar zone. Smaller articulatory constraint may thus be assumed to exist for this sound which is also frequently produced as an approximant (cf. Nicolaidis 2001, Nicolaidis & Baltazani 2011, Recasens 1991, Recasens & Pallarès 1999, Recasens & Espinosa 2007).

For the fricative /s/, analysis of the front lateral measure showed that the vowel context was significant at the 95% level with greater centrality present in the /i/ context ($F(1, 79) = 4.352$, $p < 0.041$); however, the condition by vowel interaction ($F(1, 79) = 8.734$, $p < 0.004$) showed that there was systematically more centrality in the /i/ context in the noise condition only; this was confirmed by the post-hoc tests. The vowel main effect did not reach significance at the midpoint.

For the back consonants /k, x/, the back total measure showed significantly more contact in the /i/ than /a/ context ($F(1, 159) = 7497.750$, $p < 0.0001$). Indeed, large effects were evident in the /i/ context for both consonants. Such large effects involving a distant closure target for velars as a function of the following vowel have been noted in the literature (Recasens & Espinosa 2009). Palatal production of velars in the environment of front vowels in Greek can be seen in Figure 3 (cf. Nicolaidis 2001). The palatograms show that there is extensive contact not only in the palatal region but also laterally in the alveolar zone extending up to the most anterior rows of electrodes. Vowel effects were significant and as described above at the temporal midpoint (vowel main effect: $F(1, 159) = 7014.819$, $p < 0.0001$).

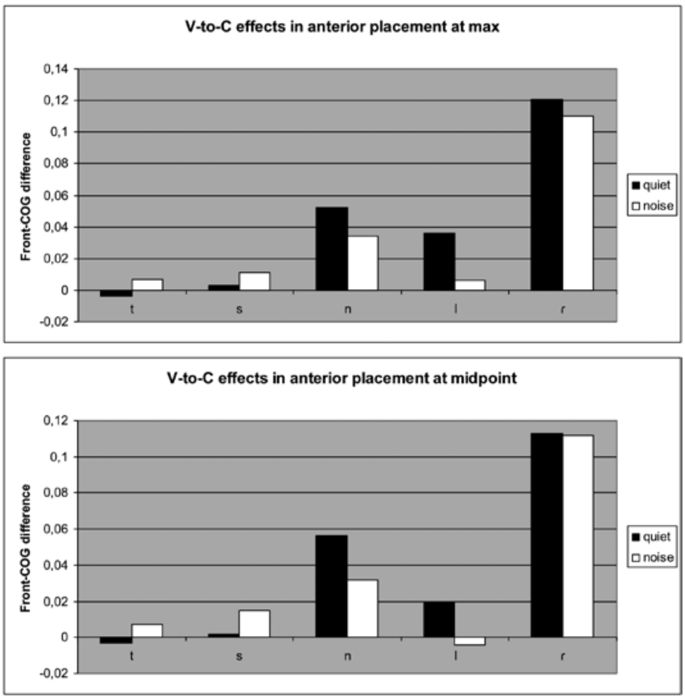


Figure 7. V-to-C effects in anterior placement of the tongue tip/blade for the consonants /t, s, n, l, ɾ/ in quiet and noise at the frame of maximum contact/constriction (upper graph) and at the temporal midpoint frame (bottom). The graphs plot the differences in the Front COG measure in the environment /i/ vs. /a/ in the two conditions.

The analysis of the back CoG for /k, x/ showed more fronted contact in the /i/ than /a/ environment (vowel main effect: $F(1, 159) = 7217.710$, $p < 0.0001$) with greater difference between environments for the fricative than the stop (consonant by vowel: $F(1, 159) = 242.049$, $p < 0.0001$). The same was evident at the temporal midpoint (vowel main effect: $F(1, 159) = 7851.660$, $p < 0.0001$; consonant by vowel: $F(1, 159) = 265.118$, $p < 0.0001$).

Finally, analysis of the back lateral measure for /x/ showed that the vowel main effect was significant ($F(1, 79) = 82.818$, $p < 0.0001$) with more centrality present in the /i/ than /a/ context for all speakers ($F(3, 79) = 27.058$, $p < 0.0001$). The same effects were evident at the midpoint (vowel main effect: $F(1, 79) = 85.879$, $p < 0.0001$; subject by vowel: $F(3, 79) = 16.346$, $p < 0.0001$).

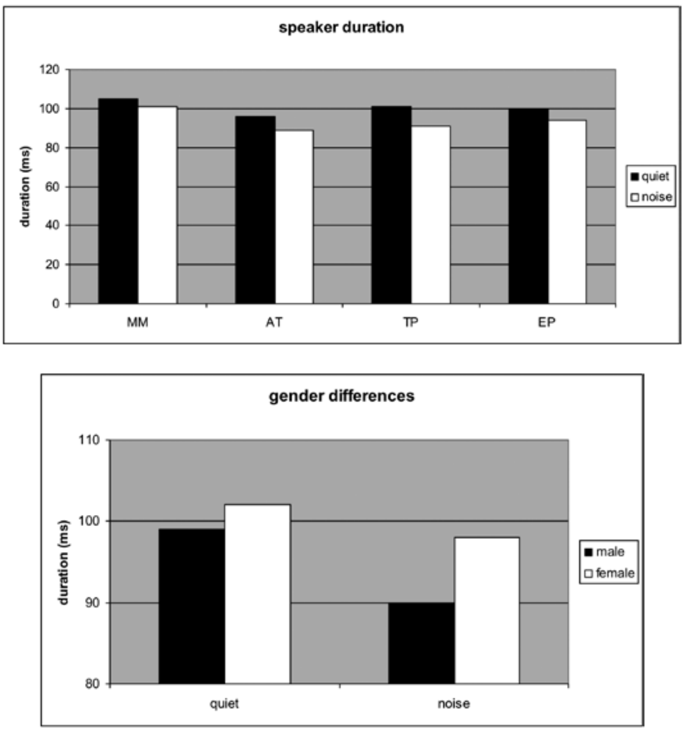


Figure 8. Average consonant duration (in ms) in the quiet and noise conditions for all subjects (upper graph), for male and female subjects (bottom).

3.4. The effect of speaker and gender

Significant speaker differences in consonantal duration were found ($F(3, 559) = 28.95$, $p < 0.0001$) with overall longest duration produced by speaker MM and the shortest by AT while the other two speakers, TP and EP, showed relatively similar durations. These differences were evident in the quiet and noise conditions ($F(3, 599) = 2.67$, $p < 0.047$). Post hoc tests showed that differences among subjects were significant but for those between TP and EP (Figure 8 and Table 2).

Longer duration (Table 2) was evident for the female speakers (gender: $F(1, 559) = 35.30$, $p < 0.0001$). These differences were significant in both conditions ($F(1, 559) = 9.20$, $p < 0.0001$) as confirmed by the post hoc tests; the means showed that greater gender differences in consonantal duration were evident in the noise condition: 8% in noise vs. 3% in the quiet condition (Figure 8). Moreover, male speakers compress duration by 9% in noise while female speakers by 4%.

Table 2. Average durations (in ms) and standard deviations pooled for all consonants in quiet and noise for the four speakers (MM, AT, TP, EP) (left) and for male and female speakers (right).

	QUIET		NOISE			QUIET		NOISE	
	AVERAGE	ST. DEV	AVERAGE	ST. DEV		AVERAGE	ST. DEV	AVERAGE	ST. DEV
MM	105	40	101	37	male	99	37	90	35
AT	96	36	89	33	female	102	40	98	36
TP	101	37	91	38					
EP	100	39	94	35					

There were also significant speaker differences in the amount of contact in the alveolar zone (front total measure) decreasing in the order EP>TP>MM>AT ($F(3, 399) = 144.198$, $p < 0.0001$).[4] These differences in contact were evident both in the noise and quiet conditions (subject by condition: ($F(3, 399) = 7.145$, $p < 0.0001$). The analysis that included gender showed more alveolar contact by the female speakers ($F(1, 399) = 62.382$, $p < 0.0001$) which was evident in the means for all consonants (consonant by gender: ($F(4, 399) = 6.271$, $p < 0.0001$) although post hoc tests showed that differences reached significance only for /s/ and /l/.

Significant speaker differences were also found in the amount of contact in the palatal region (back total measure) which decreased in the order EP>MM>TP>AT ($F(3, 399) = 176.279$, $p < 0.0001$).[5] These differences in contact were evident both in the noise and quiet conditions (subject by condition: $F(3, 399) = 21.922$, $p < 0.0001$). The analysis that included gender showed overall more palatal contact for the female speakers ($F(1, 399) = 285.224$, $p < 0.0001$) and for all consonants although post hoc tests showed that differences were significant only for /t, l, ɾ/ (consonant by gender: $F(4, 399) = 35.952$, $p < 0.0001$).

The means for the consonant by gender by vowel interaction indicated greater coarticulatory effects for the male speakers ($F(4, 399) = 8.623$, $p < 0.0001$). This was evident also at the midpoint except for /s/, which showed very small coarticulatory effects for both genders but relatively greater for the female speakers ($F(4, 399) = 13.625$, $p < 0.0001$) (Figure 9).

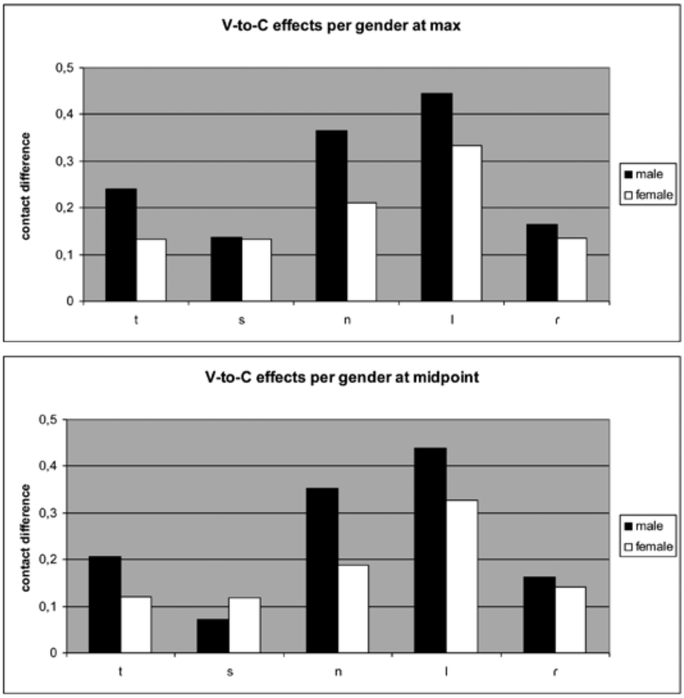


Figure 9. V-to-C effects on the tongue dorsum for the consonants /t, s, n, l, ɾ/ for the two genders at the frame of maximum contact/constriction (upper graph) and the temporal midpoint frame (bottom).

For the front CoG measure, significant speaker differences were evident ($F(3, 399) = 78.925$, $p < 0.0001$) with the male speakers showing more fronted contact compared to the female speakers (gender: $F(1, 399) = 57.128$, $p < 0.0001$); no significant speaker differences were found within gender. The means of the consonant by gender interaction ($F(4, 399) = 11.058$, $p < 0.0001$) showed more fronted contact for all consonants produced by the male speakers ($F(4, 399) = 11.058$, $p < 0.0001$), although post-hoc tests showed that significance was reached only for /ɾ/. The means of the consonant by gender by condition interaction ($F(4, 399) = 3.231$, $p < 0.013$) showed that there was an overall tendency for more fronted contact by male speakers in both conditions with only few exceptions present, i.e., /s/ in quiet and /l/ in noise.

For the fricative /s/, the analysis of the front lateral measure showed significant differences among speakers with central placement decreasing in the order MM>EP>TP>AT ($F(3, 79) = 147.994$, $p < 0.0001$). Post hoc tests showed that all differences were significant among speakers. These differences in centrality were present in both conditions. The analysis that included gender showed that more central contact was evident for the female than male speakers ($F(1, 79) = 145.134$, $p < 0.0001$) in both conditions (condition by gender: $F(1, 79) = 4.959$, $p < 0.029$); significant differences were confirmed by the post-hoc tests.

For the consonants /k, x/, significant speaker differences in the amount of contact in the palatal region (back total) were evident decreasing in the order AT>TP>MM>EP ($F(3, 159) = 29.750$, $p < 0.0001$).[6] Male speakers showed more contact than female ones ($F(1, 159) = 14.810$, $p < 0.0001$) especially so for the stop (consonant by gender: $F(1, 159) = 13.328$, $p < 0.0001$).

For the back CoG, the subject main effect showed that it decreased in the order AT>TP>EP>MM; all differences were significant. Thus, male speakers were found to have more fronted contact than female ones (gender: $F(1, 159) = 14.063$, $p < 0.0001$) while greater contextual effects were evident from the male speakers (vowel by gender: $F(1, 159) = 20.010$, $p < 0.0001$) which was also the case at the midpoint ($F(1, 159) = 22.702$, $p < 0.0001$). Finally, the analysis of the back lateral measure for /x/ showed that the subject main effect was significant; centrality decreased in the order of AT>TP>MM>EP although post hoc tests showed that not all of these differences among speakers were significant. Male speakers showed more central contact (gender: $F(1, 79) = 13.534$, $p < 0.0001$) and more context effects (vowel by gender: $F(1, 79) = 21.892$, $p < 0.0001$).

3.5. Token-to-token variability

Token-to-token variability did not differ significantly between conditions. The consonant main effect was significant showing that variability decreased in the order of l>ɾ>n>s>k>t>x ($F(6, 111) = 7.123$, $p < 0.001$).[7] Greater variability was evident in the /i/ context for all consonants but for /l/ (consonant by vowel: $F(6, 111) = 6.138$, $p < 0.001$). Token-to-token variability may be interpreted to reflect articulatory constraint and articulatory precision during production. The results show less variability (thus more articulatory constraint) for velars, the fricative /s/ and the alveolar /t/ and more variability (thus less articulatory constraint) for the liquids and nasal in agreement with the articulator constraint model proposed by Recasens et al. (1997). It should be noted that velars in the environment of /a/ show very little, relatively invariant contact. This may influence the findings reported here which do not reflect variability of the entire consonantal gesture. However, in agreement with our findings, dorsals have been reported to be highly constrained due to the involvement of the tongue dorsum in the formation of the primary constriction (Recasens et al. 1997).

4. Discussion

This study has shown that consonant production in noise is characterised by spatio-temporal adaptations. Our findings also agree with previous literature in showing increased speaker variability in Lombard speech. The discussion below will report statistically significant findings but it will also elaborate on overall tendencies which, despite not reaching significance at all times, appeared to be systematic across speakers, conditions and so on. The relatively small sample size and the increased speaker variation render the analysis of more data necessary before firm conclusions can be drawn.

In the temporal domain, a tendency for shorter durations were noted for all consonants, both voiceless and voiced, with very few exceptions present. Significant differences were however found only for /x, n/. Lu & Cooke (2008) reported slight shortening for /f/ and non alveolar plosives in English Lombard speech. Shorter consonant duration was reported by Garnier et al. (2006) for French Lombard speech while durational changes in vowels but not in consonants (for voiceless and voiced plosives) were reported for Spanish Lombard speech by Castellanos et al. (1996). Shortening of consonants was also reported for Korean loud speech (for voiceless consonants only) as well as German loud speech and French shouted speech (Kim 2005, Geumann 2001, Bonnot & Chevrie-Muller 1991). Shorter duration of intervocalic bilabial stops was reported by Schulman (1989) for loud speech which was attributed to greater velocities in the presence of greater articulatory displacements during the production of the bilabial.

The degree of shortening differed among consonants. For example, the alveolar plosive showed the smallest degree of shortening despite its relatively long segmental duration. The production characteristics of this sound which involve complete closure with lingual bracing against the palate may result in resistance to shortening. Strong articulatory (lingual and mandibular) and aerodynamic constraints involved in the production of the alveolar /s/ may also account for its small degree of shortening.

In the spatial domain, significant differences in the total amount of contact in the palatal region were found in the two conditions for /t, s, n, l, ɾ/. There was a tendency for less contact in the palatal region for all these consonants produced in noise which suggests a lowered tongue dorsum. The overall tendency observed can be interpreted in line with evidence suggesting wider jaw opening in speech conditions where greater amplitude is needed (Schulman 1989). Mooshammer et al. (2006) reported, however, significant jaw lowering for vowels in loud speech but a much less noticeable influence on consonants. In general, their study showed lower jaw position in loud speech for consonants typically produced with more open jaw positions such as /l, n/; still the condition effect was not significant even for the lateral. Significant speaker differences were however noted in jaw opening in loud speech with some speakers showing lower jaw position for /n, l/ and one speaker for /t, d/. No variation in jaw position due to vocal effort was found for sibilant sounds. Our data has also shown significant speaker differences in palatal contact across consonants and conditions.

In addition our data has shown a tendency for more contact in the alveolar region for the consonants produced in noise with just few exceptions noted. Significant differences were however evident only for /s/ and relate mainly to the articulatory patterns of one speaker. A further interesting finding is the significantly more anterior placement in the alveolar region for the consonants /n, l, ɾ/ in noise. Lack of evidence of anterior displacement in noise for /t/ may relate to the inability of EPG system used in this study to register contact in the dental area or to the fact that in Greek /t/ is classified as dental or dento-alveolar (Arvaniti 1999, Nicolaidis 1994) and thus already involves very anterior constriction bounded by the teeth. For /s/ speaker variation was evident. For /n, l, ɾ/, however, it does indicate articulation closer to the periphery of the vocal tract. Although such advanced placement could be related to increased oral pressure in the oral cavity, evidence of significantly more constricted groove width for /s/ in noise, despite the increased oral pressure, indicates possible hyperarticulation. More peripheral contact and greater amount of contact during consonant production has been interpreted to suggest stronger, hyperarticulated productions (Lavoie 2001).

Taking into consideration the overall systematic tendencies observed but also the variability noted in the statistical results reported, a tentative interpretation of the differences in consonantal production in quiet and noise will be attempted. In view of the finding that consonants are produced with more open articulation in the palatal region and somewhat shorter duration, the speaker has the following options: (a) to undershoot because the magnitude of the gesture needed for constriction in the alveolar region is larger and there is some temporal compression; (b) to compensate for the above in order to achieve the target. Reduction leading to undershoot will be guided by system-oriented constraints and the principle of economy. Indeed reduction effects on consonants frequently leading to more open articulations have been reported before for several languages including Greek (e.g., Shockey 1991, Shockey & Farnetani 1992, Nicolaidis 2001). Compensatory behaviour reflects plasticity in speech production and can occur when output purpose-oriented constraints prevail over system-oriented constraints. Severe output-oriented constraints may be assumed to exist in Lombard speech due to the need for increased intelligibility and discriminability for successful communication. Hyper-forms may thus be expected. It is known that the tongue tip/blade and dorsum can act quasi-independently. If the more open tongue dorsum position, possibly related to a lowered jaw, is considered a perturbation (Geumann et al. 1999), "a 'natural' bite block" as Schulman (1989: 310) suggests, then the tongue tip/blade can compensate for this. The speaker has again two options, i.e., to achieve the target as in normal speech produced in quiet or to hyperarticulate. The tendency for greater amount of contact in the alveolar region may suggest hyperarticulation. This can be accompanied by greater velocity of movement accounting for shorter duration. A similar finding is reported for lip movement in loud speech by Schulman (1989) where compensation for a lower jaw by the upper lip results in a closure that is more complete than in normal production. Greater lip compression for bilabial consonants produced in noise is also reported by Garnier (2008). Evidence from compensatory behaviour during consonant production in bite-block studies using electropalatography (Flege et al. 1988) also shows more linguopalatal contact in the bite block condition for English /t/ although speaker specific strategies in compensatory behaviour are also noted, e.g., more contact for one speaker during /s/ but less contact for another. However, in addition to evidence of overcompensation, the findings in the Flege et al. study showed that speakers did not always compensate completely when producing /t/ and /s/, e.g., incomplete constriction for /t/ by Arabic speakers. Similarly, findings on French /t/ produced in a bite-block condition showed speaker variability in amount of contact (more for one speaker, less for the other) suggesting speaker specific strategies in compensatory behaviour (Clairet 2008). The lack of significant results on amount of alveolar contact in the current study may relate to speaker variation for particular classes of sounds possibly stemming from speaker specific strategies in adaptive behaviour. Furthermore it relates to speaker variation in consonantal production as, for instance, the very palatalised productions of consonants in the quiet condition by one subject which were produced with increased alveolar contact in quiet and reduced contact in noise.

Still, in Lombard speech a relationship between increased lingual contact at constriction and increased oral pressure due to greater vocal effort cannot be excluded. This is also pointed out by Garnier (2008) for increased lip compression. However, her study showed increased lip compression also for the bilabial nasal suggesting that other factors may also play a role. Similarly, in the current study greater amount of contact for three speakers for the nasal /n/ was found in the noise condition. Furthermore, absence of target undershoot in our data suggests that speakers opt for a more costly form of production.

In addition to reduction, coarticulation is also driven by the preference of the speech production system to economise. Our data showed systematically smaller coarticulatory effects in tongue dorsum and tongue tip/blade activity in noise. For the tongue tip/blade, this was reflected both in context effects on place of articulation but also on the overall degree of contact in the alveolar region, although effects did not reach significance for the latter. Greater articulatory constraints on the tongue tip/blade imposed by the increased contact evident in the alveolar region may account for the reduced coarticulatory effects in line with the coarticulatory model based on articulatory constraints (Recasens et al. 1997, Recasens & Espinosa 2009). This line of interpretation cannot, however, account for the reduced coarticulatory effects in the palatal region since reduced linguo-palatal contact in this region was evident in noise. In addition, the degree of coupling between the tongue tip/blade and dorsum is expected not to be very strong for sounds such as /n, l, ɾ/, despite a tendency for increased contact in the alveolar region in noise, as their place of articulation was more anterior. The presumed more open jaw position may be assumed to play a role potentially imposing articulatory constraints although compensatory behaviour between tongue dorsum and jaw has been shown in various studies (e.g., Fletcher & Harrington 1999) while for the production of the dentoalveolars the tongue and jaw do not act coordinatively (Recasens 2002). Thus smaller coarticulatory effects of the tongue dorsum may indeed be interpreted to relate to the severe output constraints evident in Lombard speech aiming towards less contextually influenced, more 'canonical' productions for increased distinctiveness.

Despite differences in degree, the contextual effects present in the quiet and noise conditions are in line with the model of articulatory constraints. Consonants involving a high degree of articulatory constraint such as /s/ and /t/ showed small coarticulatory effects from the vowel context compared to consonants such as /l, n, ɾ/ which are not as highly constrained. The results reported in this study confirm the hypothesis of an inverse relationship between the degree of articulatory constraint and the degree of coarticulatory effects (Recasens et al. 1997, Recasens & Espinosa 2009; Farnetani 1991). Consonants produced with less coarticulatory constraint (i.e., /l, n, ɾ/), also involving smaller amounts of contact, showed larger V-to-C coarticulatory effects. The tap showed the largest V-to-C effects in anterior placement and involved the smallest amount of linguo-palatal contact. These differences in articulatory constraint among consonants were also largely confirmed by the analysis of token-to-token variability which did not show any further differences in articulatory precision between conditions. Interestingly, the relationship between articulatory constraint and coarticulatory variability is evident both in normal and Lombard speech. Thus the nature of V-to-C coarticulatory effects is similar although the degree of effects varies significantly in the two conditions. It is interesting to compare the results of coarticulatory variability and token-to-token variability with the findings of Geumann et al. (1999) who examined coarticulation and loud speech as sources of perturbation. They found that jaw position for the /s/ in German was very precise so that there was no need for compensation from the tongue tip for the formation of constriction while variability in jaw and tongue position increased from fricatives to stops to lateral and nasal sounds. Less variability near consonantal constriction, i.e., at the tongue tip/blade, and greater variability at the posterior parts of the tongue were reported.

For the /k, x/ consonants, no differences in amount of contact were evident in the two conditions. For the production of velars, very limited contact was registered by the EPG so it is not possible to examine potential condition-induced effects. Palatal production in the environment of /i/ does however involve substantial contact in the palatal region. Still, palatals were similarly produced in the quiet and noise conditions. Compared to the alveolars, this difference may relate to their retracted place of articulation and the different articulator used, the tongue dorsum. This articulator has been shown to be highly constrained during the production of dorsals (Recasens et al. 1997). Less distance also needs to be travelled by the tongue dorsum for the formation of constriction compared to the tongue tip/blade for alveolar constriction. In addition, the jaw-tongue coordination may be expected to differ in the formation of the constriction of dorsal sounds compared to alveolar sounds. Increased articulatory constraint for the velars was also confirmed by the results of token-to-token variability. In addition, smaller V-to-C effects on the stop than fricative in the horizontal, anterior/posterior, axis in the environment of /i/ vs. /a/ can be interpreted to relate to increased articulatory constraint for the formation of complete closure for the stop.

Gender differences are also evident in the data. Consonantal duration was greater for the female speakers and greater gender differences in duration in the noise condition were found. This result may be interpreted to accord with the finding that female speakers are more intelligible than male speakers in speech produced in noise (Junqua 1993). Greater distinctiveness, higher intelligibility and productions characterized by greater articulatory precision have been reported for female normal speech previously (e.g., Bradlow et al. 1996, Henton 1992, Labov 1990). Such differences have been attributed to bio-physical or social reasons (Simpson 2001, 2003, 2009; Labov 1990; Henton, 1995). Overall, greater vowel durations and utterance durations by female speakers have been reported previously for normal speech for several languages (e.g., Ericsdotter & Ericsson 2001, Simpson 2001, Simpson 2002, Byrd 1992, Whiteside 1996) although contradictory results have also been reported (e.g., Simpson 2001, 2002, Simpson & Ericsdotter 2003). Simpson & Ericsdotter (2003) report greater stressed vowel duration for female speakers but greater consonantal durations for male speakers in their study suggesting that clarity may be restricted to places of prominence. The current study has provided evidence of longer consonantal durations by female speakers in Greek Lombard speech. Interesting production differences in articulatory placement and degree of contact were also found between genders. Male speakers had overall more anterior placement in the alveolar region for /t, s, n, l, ɾ/ and the palatal region for /k, x/. They also produced /t, s, n, l, ɾ/ with overall less tongue-palate contact in both the alveolar and palatal regions and /k, x/ with more palatal contact. Thus, for the female speakers there was more extensive contact of the tongue tip/blade and dorsum with the hard palate during the production of /t, s, n, l, ɾ/. The fricative /s/ was also produced with increased centrality by female speakers. Such findings may relate to the smaller V-to-C coarticulatory effects evident for female speakers. However, smaller coarticulatory effects were also evident on the dorsals for the female speakers despite reduced palatal contact compared to the male speakers. Although anatomical/mechanical constraints and gender-related aerodynamic differences may play a role, the results indicate that such spatio-temporal patterns may also function to enhance the intelligibility of female speakers in both quiet and noise conditions.

Overall durational, articulatory and coarticulatory differences between speech produced in normal vs. noise conditions have been reported in this study. There are several limitations in this study that need to be considered in future research: (a) the relatively small sample size; more data is needed for more robust statistical results and interpretations; (b) the type of material analysed; less controlled material in communicative situations needs to be used in order to examine communicatively driven modifications; it is interesting that even with the controlled material examined in this study, where the level of communicative involvement is low or absent, the effects identified point towards strategies used for maximal efficiency in communicative situations in adverse conditions; (c) different levels and types of noise can also be used to examine potential differentiation on the nature and degree of effects; (d) analysis of sentence duration and in general rate differences in quiet vs. noise need to be examined; (e) more subjects also need to be studied for more robust gender related generalisations to be made; (f) anatomical differences in palatal shape need to be considered in the interpretation of results. Some work towards the above is currently underway.

5. Conclusions

The results of this study have shown that there are important spatio-temporal modifications in Lombard speech and presence of variability. Although increased vocal effort may relate to some of the patterns found, there is some first evidence that speakers use adaptive strategies to compensate for the severe output constraints present in Lombard speech due to the need for increased intelligibility and discriminability for successful communication, in line with the framework of adaptive variability proposed by the H&H model. Results show that consonants are produced with more open articulation in the palatal region and somewhat shorter duration in the presence of noise. The speaker has thus the following options: (a) to undershoot (system oriented behaviour), (b) to compensate in order to achieve the target (output oriented behaviour). The tendency for greater amount of contact in the alveolar region, more anterior placement for selected consonants and more constricted groove width for /s/ may suggest hyperarticulation. This can be accompanied by greater velocity of movement due to greater articulatory displacement accounting for shorter duration.

Evidence of lack of undershoot, production of hyperarticulated forms and reduced coarticulation points towards speech modifications that can be communicatively driven.

Address of the Author

Katerina Nicolaidis, Department of Theoretical and Applied Linguistics, School of English, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece < knicol@enl.auth.gr>

Notes

* This paper is written in memory of Edda Farnetani, a colleague and friend, who worked extensively on theoretical issues included in the current study and with the experimental methodology used. Her memory and her insightful and inspiring work will always be with us.
I would like to thank the editors of the Journal for the invitation to contribute to this issue. Special thanks are due to Prof. George Papanikolaou, Director of the Laboratory of Electroacoustics & TV Systems of the Department of Electrical and Computer Engineering at Aristotle University, for making available the lab, audio and calibration equipment for the recording. His assistance in processing the multi-talker babble noise file and in making sure that the recording was methodologically accurate is greatly appreciated. Many thanks are also due to Christos Zarras and Kostis Konstadoudakis for assisting with the recording. I am also very grateful to Martin Cooke for providing the multi-talker babble noise file. Special thanks are also due to Tasos Paschalis for technical assistance, Alan Wrench for hardware and software support, Christos Nakas and Anna Sfakianaki for assistance with the statistical analysis, and to all the subjects of the study.

[1] Note that the global total and global CoG were calculated only for the palatal production of /k, x/ in the environment of /i/.
[2] More contact was found in noise for /t/ and /ɾ/ (for the latter in the /i/ context only) by speaker MM, and for /s/ by EP.
[3] I.e., /s/, /ɾ/ were less fronted for EP, and /n/, /l/ for AT in noise.
[4] All differences were significant but between TP, MM.
[5] All differences were significant but between EP and MM.
[6] All differences were significant but between AT, TP and TP, MM.
[7] Pairwise comparisons showed that /l/ and /ɾ/ differed significantly from /s, k, t, x/, /k/ from /x/, and /x/ from /n/.

Bibliographical references

AMAZI Deborah K. & Sharon R. GARBER 1982. The Lombard sign as a function of age and task. *Journal of Speech and Hearing Research* 25. 581-585.

ANGLADE Yolande & Jean-Claude JUNQUA 1990. Acoustic-phonetic study of Lombard speech in the case of isolated words. In TORRES Lius, Enrique

MASGRAU & Miguel A. LAGUNAS, (eds.). *Signal Processing V: Theories and Applications*. Amsterdam: Elsevier. 1195-1198.

Articulate Assistant User Guide. Version 1.18. 2003-2008. Articulate Instruments Ltd. www.articulalateinstruments.com

ARVANITI Amalia 1999. Standard Modern Greek. *Journal of the International Phonetic Association* 29. 167-172.

ARVANITI Amalia 2007. Greek Phonetics: The State of the Art. *Journal of Greek Linguistics* 8. 97-208.

BOND Z.S., Thomas J. MOORE & Beverley GABLE 1989. Acoustic-phonetic characteristics of speech produced in noise and while wearing an oxygen mask. *Journal of the Acoustical Society of America* 85. 907-912.

BONNOT Jean-François & Claude CHEVRIE-MULLER 1991. Some effects of shouted and whispered conditions on temporal organisation. *Journal of Phonetics* 19. 473-483.

BRADLOW Ann R., Gina M. TORRETTA & David. B. PISONI 1996. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20. 255-272.

BYRD Dani 1992. Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of the Acoustical Society of America* 92. 593-596.

CASTELLANOS Antonio, José-Miguel BENEDÍ & Francisco CASACUBERTA 1996. An analysis of general acoustic-phonetic features for Spanish speech produced with the Lombard effect. *Speech Communication* 20. 23-35.

CLAIRET Sandrine 2008. Compensating for a bite bock in /t/ and /d/ in French: electropalatography study. In SOCK Rudolph, Susanne FUCHS and Yves LAPRIE (eds.). *Proceedings of ISSP'08. VIIIth International Seminar on Speech Production* (December 8-12, 2008). Strasbourg. France. 413-416.

DAVIS Chris, Jeesun KIM, Katja GRAUWINKEL & Hansjörg MIXDORFF 2006. Lombard speech: Auditory (A), Visual (V) and AV effects. In HOFFMANN Rüdiger & Hansjörg MIXDORFF (eds.). *Proceedings of the IIIrd International Conference of Speech Prosody*. Dresden, Germany: TUD Press. 248-252.

DREHER John J. & John J. O'NEILL 1957. Effects of ambient noise on speaker intelligibility for words and phrases. *Journal of the Acoustical Society of America* 29. 1320-1323.

ERICSDOTTER Christine & Anna M. ERICSSON 2001. Gender differences in vowel duration in read Swedish: Preliminary results. In *Proceedings of Fonetik 2001. XIVth Swedish Phonetics Conference* (May 30 - June 1, 2001). *Working Papers, Department of Linguistics, Lund University* 49. 34-37.

FARNETANI Edda 1991. V-C-V lingual coarticulation and its spatiotemporal domain. In HARDCASTLE William J. & Alan MARCHAL (eds.). *Speech Production and Speech Modelling*. Dordrecht: Kluwer Academic Publishers. 93-130.

FARNETANI Edda & Andrea PROVAGLIO 1991. Assessing variability of lingual consonants in Italian. *Quaderni del Centro di Studio per le Ricerche di Fonetica* 10. 117-145.

FLEGE James E., Samuel G. FLETCHER & Abdullah HOMIEDAN 1988. Compensating for a bite block in /s/ and /t/ production: palatographic, acoustic, and perceptual data. *Journal of the Acoustical Society of America* 83. 212-228.

FLETCHER Janet & Jonathan HARRINGTON 1999. Lip and jaw coarticulation. In HARDCASTLE William J. & Nigel HEWLETT (eds.). *Coarticulation. Theory, Data and Techniques.* Cambridge: Cambridge University Press. 164-175.

GARNIER Maëva 2008. May speech modifications in noise contribute to enhance audio-visible cues to segment perception? In GÖCKE Ronald, Patrick LUCEY & Simon LUCEY (eds.). *Proceedings of AVSP '08. International Conference on Auditory-Visual Speech Processing* (September 26-29, 2008). Moreton Island, Queensland. Australia. 95-100.

GARNIER Maëva, Lucie BAILLY, Marion DOHEN, Pauline WELBY & Hélène LŒVENBRUCK 2006. An acoustic and articulatory study of Lombard speech: global effects on the utterance. In *Proceedings of Interspeech 2006. IXth International Conference on Spoken Language Processing* (September 17-21, 2006). Pittsburgh. Pennsylvania. 17-22.

GARNIER Maëva, Marion DOHEN, Hélène LŒVENBRUCK, Pauline WELBY & Lucie BAILLY 2006. The Lombard effect: a physiological reflex or a controlled intelligibility enhancement? In YEHIA Hani Camille, Didier DEMOLIN & Rafael LABOISSIERE (eds.). *Proceedings of ISSP7. VIIth International Seminar on Speech Production*. Belo Horizonte, Brazil. 255-262.

GARNIER Maëva, Nathalie HENRICH & Danièle DUBOIS 2010. Influence of sound immersion and communicative interaction on the Lombard effect. *Journal of Speech, Language and Hearing Research* 53. 588-608.

GEUMANN Anja 2001. *Invariance and variability in articulation and acoustics of natural perturbed speech*. University of Munich. PhD Dissertation.

GEUMANN Anja, Christian KROOS & Hans G. TILLMAN 1999. Are there compensatory effects in natural speech?. In OHALA John J., Yoko HASEGAWA, Manjari OHALA, Daniel GRANVILLE & Ashlee C. BAILEY (eds.). *Proceedings of the XIVth International Congress of Phonetic Sciences* (August 1-7, 1999). San Francisco, California. Vol. 1. 399-402.

GIBBON Fiona & Katerina NICOLAIDIS 1999. Palatography. In HARDCASTLE William J. & Nigel HEWLETT (eds.). *Coarticulation. Theory, Data and Techniques*. Cambridge: Cambridge University Press. 229-245.

HENTON Caroline 1992. The abnormality of male speech. In WOLF George (ed.). *New Departures in Linguistics.* New York: Garland Publishing Co. 27-59.

HENTON Caroline. 1995. Cross-language variation in the vowels of female and male speakers. In ELENIUS Kjell & Peter BRANDERUD (eds.). *Proceedings of the XIIIth International Congress of Phonetic Sciences* (August 13-19, 1995). Stockholm. Sweden. Vol. 4. 420-423.

JUNQUA Jean-Claude 1993. The Lombard reflex and its role on human listeners and automatic speech recognisers. *Journal of the Acoustical Society of America* 93. 510-524.

JUNQUA Jean-Claude 1994. A duration study of speech vowels produced in noise. In *Proceedings of ICSLP 94. IIIrd International Conference on Spoken Language Processing* (September 18-22, 1994). Yokohama, Japan. 419-422.

JUNQUA Jean-Claude 1996. The influence of acoustics on speech production: a noise induced stress phenomenon known as the Lombard reflex. *Speech Communication* 20. 13-22.

JUNQUA Jean-Claude, Steven FINCKLE & Ken FIELD 1999. The Lombard effect: a reflex to better communicate with others in noise. In *Proceedings of ICASSP '99. IEEE International Conference on Acoustics, Speech, and Signal Processing* (March 15-19, 1999). Phoenix, Arizona. Vol. 4. 2083-2086.

KIM Sunhee 2005. Durational characteristics of Korean Lombard Speech. In *Proceedings of Interspeech 2005 – Eurospeech. IXth European Conference on Speech Communication and Technology* (September 4-8, 2005). Lisbon, Portugal. 2901-2904.

LABOV William 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2. 205-254.

LANE Harlan & Bernard TRANEL 1971. The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research* 14. 677-709.

LAVOIE Lisa M. 2001. *Consonant strength: phonological patterns and phonetic manifestations*. New York: Garland Publishing.

LETOWSKI Tomasz, Tom FRANK & Jane CARAVELLA 1993. Acoustical properties of speech produced in noise presented through supra-aural earphones. *Ear & Hearing* 14. 332-338.

LINDBLOM Björn. 1990. Explaining phonetic variation: a sketch of the H&H theory. In HARDCASTLE William J. & Alan MARCHAL (eds.). *Speech Production and Speech Modelling*. Dordrecht: Kluwer Academic Publishers. 403-439.

LOMBARD Étienne 1911. Le signe de l'élévation de la voix. *Annales des Maladies de l'Oreille, du Larynx du Nez et du Pharynx* 37. 101-119.

LU Youyi & Martin COOKE 2008. Speech production modifications produced by competing talkers, babble, and stationary noise. *Journal of the Acoustical Society of America* 124. 3261-3275.

MOOSHAMMER Christine, Philip HOOLE & Anja GEUMANN 2006. Interarticulator cohesion within coronal consonant production. *Journal of the Acoustical Society of America* 120. 1028-1039.

NICOLAIDIS Katerina 1994. Aspects of lingual articulation in Greek: an electropalatographic study. In PHILIPPAKI-WARBURTON Irene, Katerina NICOLAIDIS & Maria SIFIANOU (eds.). *Themes in Greek Linguistics*. Amsterdam: John Benjamins. 225-232.

NICOLAIDIS Katerina 1999. An electropalatographic study of Vowel-to-Consonant and Consonant-to-Vowel coarticulation. In *Greek Linguistics '97. IIIrd International Conference on the Greek Language*. Athens. Greece. 400-408.

NICOLAIDIS Katerina 2001. An electropalatographic study of Greek spontaneous speech. *Journal of the International Phonetic Association* 31. 67-85.

NICOLAIDIS Katerina 2002. Durational variability in Vowel-Consonant-Vowel sequences in Greek: The influence of phonetic identity, context and speaker. In MAKRI-TSILIPAKOU Marianthi (ed.), *Selected Papers on Theoretical and Applied Linguistics from the XIVth International Symposium* (April 20-22, 2000). Aristotle University of Thessaloniki. 280-294.

NICOLAIDIS, Katerina & Mary BALTAZANI 2011. An electropalatographic and acoustic study of the Greek rhotic in /Cr/ clusters. In LEE Wai-Sum

& Eric Zee (eds.). *Proceedings of the XVII International Congress of Phonetic Sciences* (August 17-21, 2011). Hong Kong, China. 1474-1477.

Nicolaidis Katerina & Raphael Rispoli 2005. The effect of noise on speech production: An acoustic study. *Studies in Greek Linguistics* 25. 415-426.

Patel Rupal & Kevin W. Schell 2008. The influence of linguistic content on the Lombard effect. *Journal of Speech, Language and Hearing Research* 51. 209-220.

Pisoni David B., Robert H. Bernacki, Howard C. Nusbaum & Moshe Yuchtman 1985. Some acoustic-phonetic correlates of speech produced in noise. In *Proceedings of ICASSP '85. IEEE- International Conference on Acoustics, Speech, and Signal Processing*. 1581-1584.

Pittman Andrea L. & Terry L. Wiley 2001. Recognition of speech produced in noise. *Journal of Speech, Language and Hearing Research* 44. 487-496.

Ramez R., 1992. Changes in speaker articulation due to ambient noise. CRIN-CNRS/INRIA Lorraine, ESPRIT II(5516) ROARS Project, D25 Report, September 1992.

Recasens Daniel 1991. On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics* 19. 267-280.

Recasens Daniel 2002. Patterns of jaw coarticulatory direction and linguomandibular coordination in VCV sequences. In Hall Tracy A. (ed.). *Papers on Phonetics and Phonology: The Articulation, Acoustics and Perception of Consonants* (ZAS Papers in Linguistics). Berlin: Zentrum für Allgemeine Sprachwissenschaft. 83-92.

Recasens Daniel & Aina Espinosa 2007. Phonetic typology and positional allophones for alveolar rhotics in Catalan. *Phonetica* 63. 1-28.

Recasens Daniel & Aina Espinosa 2009. An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America* 125. 2288-2298.

Recasens Daniel, Edda Farnetani, Jordi Fontdevila & Maria Dolors Pallarès 1993. An electropalatographic study of alveolar and palatal consonants in Catalan and Italian. *Language and Speech* 36. 213-234.

Recasens Daniel & Maria Dolors Pallarès 1999. The study of /ɾ/ and /r/ in the light of the 'DAC' coarticulation model. *Journal of Phonetics* 27. 143-169.

Recasens Daniel, Maria Dolors Pallarès & Jordi Fontdevila 1997. A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*. 102. 544-561.

Schulman Richard 1989. Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America* 85. 295-312.

Shockey Linda 1991. Electropalatography of conversational speech. In *Proceedings of the XIIth International Congress of Phonetic Sciences* (August 19-24, 1991). Aix-en-Provence, France. Vol. 3. 10-13.

Shockey Linda & Edda Farnetani 1992. Spontaneous speech in English and Italian. In *Proceedings of ICSLP 92. IInd International Conference on Spoken Language Processing* (October 13-16, 1992). Banff, Canada. 767-770.

Simpson Adrian P. 2001. Dynamic consequences of differences in male and female vocal tract dimensions. *Journal of the Acoustical Society of America* 109. 2153-2164.

SIMPSON Adrian P. 2002. Gender-specific articulatory-acoustic relations in vowel sequences. *Journal of Phonetics* 30. 417-435.

SIMPSON Adrian P. 2003. Possible articulatory reasons for sex-specific differences in vowel duration. In *Proceedings of ISSP 6. VIth International Seminar on Speech Production* (December 7-10, 2003). Sydney, Australia. 261-266.

SIMPSON Adrian P. 2009. Phonetic differences between male and female speech. *Language and Linguistics Compass* 3. 621-640.

SIMPSON Adrian P. & Christine ERICSDOTTER 2003. Sex-specific durational differences in English and Swedish. In SOLÉ Maria J., Daniel RECASENS & Joaquim ROMERO (eds.). *Proceedings of the XVth International Congress of Phonetic Sciences* (August 3-9, 2003). Barcelona, Spain. 1113-1116.

STEENEKEN Herman J.M. & John H.L. HANSEN 1999. Speech under stress conditions: overview of the effect on speech production and on system performance. In *Proceedings of ICASSP '99. IEEE International Conference on Acoustics, Speech, and Signal Processing* (March 15-19, 1999). Phoenix, Arizona. Vol. 4. 2079-2082.

TAKIZAWA Yumi & Masahiro HAMADA 1990. Lombard speech recognition by formant-frequency-shifted LPC cepstrum. In *ICSLP 90. Ist International Conference on Spoken Language Processing* (November 18-22, 1990). Kobe, Japan. 293-296.

VAN SUMMERS W., David B. PISONI, Robert H. BERNACKI, Robert I. PEDLOW & Michael A. STOKES 1988. Effects of noise on speech production: acoustic and perceptual analyses. *Journal of the Acoustical Society of America* 84. 917-928.

VAN SUMMERS W., Keith JOHNSON, David B. PISONI & Robert H. BERNACKI 1989. An addendum to 'Effects of noise on speech production: acoustic and perceptual analyses' [J. Acoust.Soc.Am.84, 917-928 (1988)]. *Journal of the Acoustical Society of America* 86. 1717-1721.

VARADARAJAN Vaishnevi S. & John H.L. HANSEN 2006. Analysis of the Lombard effect under different types and levels of noise with application to in-set speaker ID system. In *Proceedings of Interspeech 2006. IXth International Conference on Spoken Language Processing* (September 17-21, 2006). Pittsburgh. Pennsylvania. 937-940.

WELBY Pauline 2006. Intonational differences in Lombard speech: looking beyond F0 range. In Rüdiger HOFFMANN & Hansjörg MIXDORFF (eds.). *Proceedings of the IIIrd International Conference on Speech Prosody*. Dresden, Germany: TUD Press. 763-766.

WHITESIDE Sandra P. 1996. Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences. *Journal of the International Phonetic Association* 26. 23-40.