



INTERACTION OF INTRINSIC VOWEL AND CONSONANT DURATIONAL CORRELATES WITH FOREIGNER DIRECTED SPEECH

JOANNA SANKOWSKA^a, MARÍA LUISA GARCÍA LECUMBERRI^b
AND MARTIN COOKE^{a,b,c}

^{a,b}*Universidad del País Vasco* ^c*Ikerbasque (Basque Science Foundation)*
^a*j.sankowska@hotmail.com* ^b*garcia.lecumberrri@ehu.es* ^c*m.cooke@ikerbasque.org*

ABSTRACT

Foreigner-directed and Lombard speech are two examples of speech modes that have increased intelligibility compared to normal speech. Investigating the interactions between altered speech modes and phonological contrasts may throw light on the question of which details are vital in intelligibility enhancement. The present study compares the production of vowel shortening in English, a duration-based voicing correlate, realized in a listener-directed speech style (foreigner-directed speech) with native adult-directed speech and another listener-directed speech mode (Lombard speech). British speakers completed a communicative task in cooperation with an adult native speaker or adult foreigner, in quiet and in noisy conditions. Speaker productions were analyzed to examine the changes in the duration of the target vowels and following plosive consonants.

The results show that vowel shortening was present in the three speech styles. The durational voicing correlate was maintained in foreigner-directed and reduced in Lombard speech when compared with native adult-directed speech. Consonant durational differences were enhanced in foreigner-directed but reduced in Lombard speech relative to native adult-directed speech. The results suggest that foreigner-directed speech may be more intelligible in quiet conditions than Lombard speech, but less when both are presented with the same amount of noise.

KEYWORDS: Foreigner-directed speech; Lombard speech.

1. Introduction

Speech has been shown to accommodate to meet the needs of the listener. For instance, people tend to speak more loudly in the presence of noise (Lombard 1911; Dreher and O'Neill 1957; Summers et al. 1988; Garnier et al. 2006), and

more slowly and with increased pause duration when addressing a child (DePaulo and Coleman 1986; Biersack et al. 2005). A further listener-directed speech style is foreigner-directed speech (FDS), a speech mode that aims to enhance intelligibility. It is addressed to adult interlocutors with perceived limited linguistic capacity e.g. foreign learners. Previous research devoted to FDS has found a decrease in speech rate (Biersack et al. 2005; Scarborough et al. 2007), an expansion in vowel space (Knoll et al. 2004; Uther et al. 2007), and an increase in the duration of vowels (Scarborough et al. 2007; Ashby 2004). FDS can be compared with other speech styles that are modified in order to meet the communicative needs of the target audience. One of those listener-directed speech styles is Lombard speech (LS) that describes alterations in speaker vocal production in noisy environments. Previous research has found measurable differences in duration, pitch, intensity, and formant frequencies in the presence of noise (Summers et al. 1988; Junqua 1993). Lombard speech is typically more intelligible than speech produced in quiet when both are presented in equivalent amounts of noise (Dreher and O'Neill 1957; Summers et al. 1988; Lu and Cooke 2008). In Lombard speech, reported differences in duration relate to the increase in the duration of words (Summers et al. 1988), syllables (Patel and Schell 2008), certain consonants (Lu 2010) and vowels (Junqua 1993; Ashby 2004). Junqua also reports a slight decrease in the duration of consonants and Lu reports the decrease in the duration of voiceless labiodental fricatives and non-alveolar plosives.

An intriguing possibility is that changes in speech production induced by listeners' needs might interact with phonetic correlates observed in normal speech. If foreigner-directed and Lombard speech aim to increase intelligibility then some cues may also be enhanced in order to emphasize contrasts. One well-known duration-based feature of English phonology is vowel shortening, which describes the phenomenon where a vowel followed by a voiceless consonant in the same syllable is shorter than it would be when followed by a voiced consonant (Wells 1981; Gimson 1989; Cruttenden 2001) e.g. the vowel /i:/ in the word *beat* is shorter than in the word *bead*. The current study seeks to describe possible interactions between duration-based contrasts and the properties of speech directed to a non-native speaker in real interactions. We investigate if vowel shortening is enhanced, maintained or reduced in FDS. Furthermore, we compare the changes occurring in FDS with those that are present in LS.

2. Method

Changes in vowel and consonant (plosive) duration were investigated in a communicative task in three different conditions: (i) in quiet when interacting

with a native adult interlocutor, (ii) in quiet when interacting with a foreign adult interlocutor, and (iii) in the presence of stationary noise when interacting with a native adult interlocutor. Speakers produced target words in frame sentences which varied in content but were syntactically equivalent. Speaker productions were analyzed to examine differences in the duration of target vowels and following plosive consonants.

2.1. Task

Pairs of participants were involved in a communicative task designed for the purpose of this study. During the task, one of the participants (the speaker) had to compose sentences such as *Mr Gar Ven will say dog to Mr Garve again* using words on sets of cards prepared by the investigator. The sets of cards were attached to the table so that the subject could not change the structure of the sentence. Further, each set had a number in order to avoid mixing the words from two different sets. The speaker was instructed to produce the sentence for the other participant (the listener) to find it on the list of all possible sentences. The listener was instructed to find the sentence and repeat it for confirmation. After completing the task the participants were instructed to swap roles and complete the task again using different sets of cards. This procedure was the same for all participants. However, the foreigner's speech material was not included in the analysis since their role was to elicit FDS while taking part in a communicative task with a native interlocutor.

2.2. Materials

Target words were chosen which contained long vowels /i:/, /a:/, /ɔ:/ or short vowels /ɪ/, /æ/, /ɒ/, followed by either voiceless plosives /p/, /t/, /k/ or voiced plosives /b/, /d/, /g/. The target words were minimal pairs of real English words with the structure CVC and appeared sentence-medially, e.g.:

- (1a) Mr Ghee Van will say *ford* to Mr Pork again,
- (1b) Mr Key Fan will say *fort* to Mr Borg again,
- (1c) Mr Gar Den will say *bag* to Mr Piece again,
- (1d) Mr Car Ten will say *back* to Mr Bees again.

The first names that occurred sentence-initially were chosen to allow investigation of vowel shortening on the word boundary. These data and results are not

presented in this paper. Vowels and consonants in coda position in the target words were used for the analysis. There were 432 tokens collected in total (144 in each of the three conditions).

2.3. Procedure

Each participant sat in a sound-attenuating booth and produced the sentences for the other person sitting in front of them but separated by an opaque screen built for the purpose of this study, meant to prevent the use of visual cues in the task. Each participant took part in 3 sessions which together lasted around 1 hour, including a prior practice and breaks. In the first session, speech was recorded in quiet conditions. During the second session both subjects were exposed to speech shaped noise at 85 dB SPL delivered through headphones. The third session involved completing the task in cooperation with an adult foreign interlocutor in quiet conditions. Recordings were done using a MOTU 8pre FireWire audio interface with Audio Desk 2 software and head-mounted microphones (Sennheiser MZA 900 P). The order of sessions was counterbalanced across speakers.

2.4. Speakers

Four native speakers of Standard British English with comparable southern English accents were recruited from the students and staff of the Computer Science Department at the University of Sheffield, UK. They were recorded in interactions with a British adult interlocutor and a foreign (Chinese) adult interlocutor with noticeable foreign (Mandarin) accent. None of the participants had any known history of speech or hearing impairment.

3. Results

All acoustic feature measurements were obtained using PRAAT (Boersma and Weenink 2005). First, the durations of vowels and following plosives in the target words were measured. Second, the vowel durational correlate was calculated as the difference between the vowel duration in the voiced context and the vowel duration in the voiceless context, and the consonant durational contrast was calculated as the difference between the duration of the voiced consonant and the duration of the voiceless consonant.

3.1. Vowel shortening

Before analyzing the durational contrasts we checked that vowel shortening was present in all speech styles. Figure 1 shows mean vowel and consonant duration in the adult-directed speech (ADS), foreigner-directed speech (FDS), and Lombard speech (LS) conditions. Analysis of vowel and plosive duration revealed that vowel shortening due to the voicing properties of the following plosive was observed for both long and short vowels in ADS. The duration of long and short vowels was increased ($t(70) = 4.083$, $p < .001$ and $t(70) = 4.150$, $p < .001$ respectively) when followed by a voiced plosive. Analysis of the consonants also confirmed durational differences. Voiced plosives that followed both long and short vowels were shorter than voiceless ones ($t(70) = 4.277$, $p < .001$ and $t(70) = 5.882$, $p < .001$ respectively). Similar results were found for vowel and plosive duration for foreigner-directed speech. Long and short vowels were significantly longer ($t(70) = 3.833$, $p < .001$ and $t(70) = 4.754$, $p < .001$ respectively) when followed by a voiced plosive. Analysis of consonant duration also showed that voiced plosives that followed both long and short vowels were shorter than voiceless ones ($t(70) = 8.113$, $p < .001$ and $t(70) = 8.417$, $p < .001$ respectively). These results suggest that vowel shortening was still observed in this altered speech style. However, for Lombard speech, this pattern of results was observed only for vowels. The duration of long vowels was increased when the vowel was followed by a voiced plosive as opposed to a voiceless one ($t(70) = 2.544$, $p < .05$) and the same trend was found for the short vowels ($t(70) = 4.216$, $p < .001$). However, the analysis of the consonant duration showed differences between the plosives following long vowels and those following short vowels. For the voiced plosives following long vowels, there was a tendency for the duration to be decreased but the result fell short of significance ($t(70) = 1.916$, $p = .059$). This may be due to the inter-speaker variability and the fact that some speakers may use other strategies to signal vowel shortening. However, the duration of the voiced plosives following short vowels was significantly shorter than voiceless plosives ($t(70) = 2.103$, $p < .05$).

3.2. Vowel-duration-based voicing correlate

Durational differences for vowels followed by voiceless vs. voiced consonants were analyzed by a repeated measures ANOVA with two factors of Length and Style, Length having two levels (intrinsically long vs. intrinsically short vowels), and Style having three levels (ADS vs. LS vs. FDS).

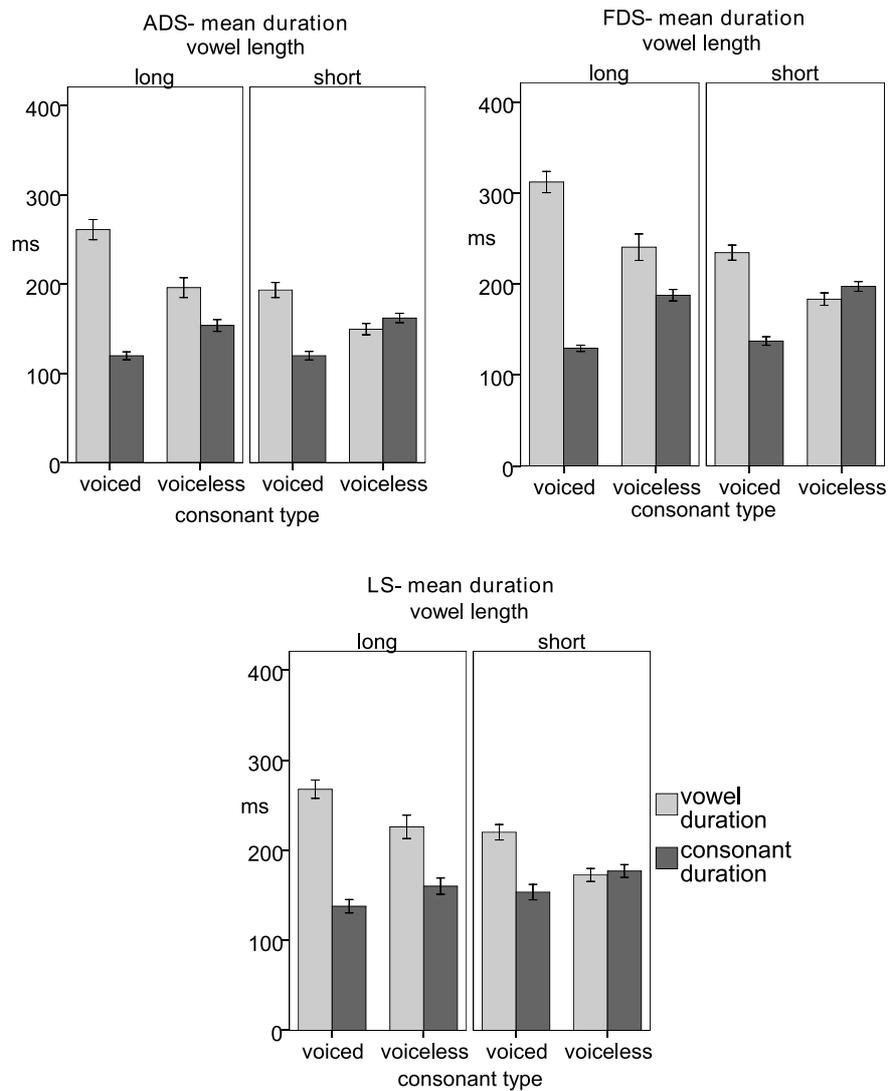


Figure 1. Mean vowel and plusive duration (error bars: +/- 1 Standard Error).

Figure 2 shows mean vowel duration-based voicing contrasts in the three speaking styles. We found no difference between the styles for short vowels. However, for long vowels we found that the durational contrast was decreased

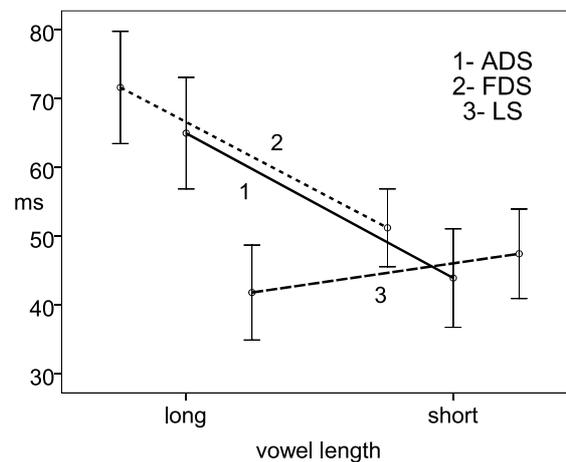


Figure 2. Mean vowel durational contrast.

for LS relative to ADS and FDS. What is more, the results showed that in ADS and FDS the durational contrast was smaller for short vowels as opposed to long vowels and there was no difference in LS. The ANOVA analysis of vowel duration-based voicing contrasts confirmed these impressions and indicated a significant interaction between Style and Length ($F(2,34) = 5.51, p < .01$) and a significant effect of Style ($F(2,34) = 6.612, p < .01$). Different tendencies were observed for long and short vowels. For the long vowels, the durational contrast was significantly enhanced for FDS relative to LS ($p < .001$). It was significantly reduced for LS relative to ADS ($p < .01$), and no difference was found between ADS and FDS. As for the short vowels, no difference was found between the conditions. These results indicate that, at least for the long vowels, duration-based contrasts are in fact maintained in foreigner-directed speech and reduced in Lombard speech if both are compared to ADS as a baseline. Further analysis of the data revealed that, in ADS, the duration-based voicing contrast was significantly smaller for short vowels than it was for long vowels ($p < .05$). There was a tendency for the durational contrast to be smaller for short vowels than for long vowels in FDS but the difference fell short of significance ($p = .056$). No difference was found in LS between the long and short vowels. A comparison of adult-directed and Lombard speech showed that there was no significant interaction but a significant effect of Length $F(1,35) = 5.4, p < .05$. The duration-based voicing contrast was reduced for short vowels rather than long vowels.

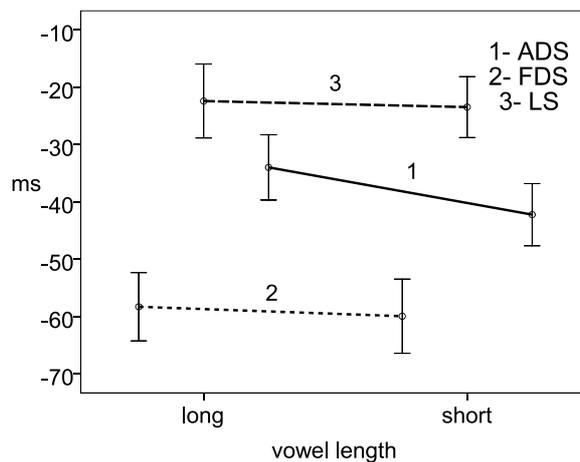


Figure 3: Mean consonant durational contrasts.

3.3. Duration-based consonant voicing contrast

Figure 3 shows mean durational differences between voiceless and voiced consonants in the three speech styles. Consonant durational differences were enhanced for FDS and reduced for LS relative to ADS. A repeated measures ANOVA highlighted a significant effect of Style ($F(2,34) = 28.069$, $p < .001$). The durational difference between voiced and voiceless consonants was enhanced for FDS relative to both ADS ($p < .01$) and LS ($p < .0001$). We also found that the durational difference was decreased for LS relative to ADS ($p < .01$). There was no statistically significant effect of Length. These results indicate that, although FDS and LS are both listener-directed speech styles, they exhibit different patterns of altering the speech. It seems that for plosives the duration-based voicing contrasts are reduced in LS but enhanced in FDS relative to ADS.

4. Discussion

The aim of this study was to investigate a duration-based phonological contrast in speech directed to non-native speakers while conducting a communicative task, and compare it with adult-directed and Lombard speech. First, we investigated if vowel shortening before voiceless consonants was present in all speech

styles. We found that speakers produced shorter vowels when followed by voiceless plosives rather than by voiced ones. Also, voiced plosives were shorter than their voiceless counterparts in all tested speaking styles (although in the case of Lombard speech for voiced plosives following long vowels the shortening did not quite reach statistical significance). This suggests that although foreigner-directed and Lombard speech are altered modes of speech, vowel shortening is still present. Second, we investigated whether durational contrasts were maintained, reduced or enhanced in the listener-directed speech styles. Analysis of the FDS data showed that vowel duration-based consonant voicing contrasts are in fact preserved in foreigner-directed speech. Different tendencies are present in Lombard speech, i.e. vowel duration-based contrasts are reduced in Lombard speech if compared to ADS as a baseline. What is more, at least for the long vowels, the durational contrast was significantly enhanced for FDS when compared with LS. Further analysis revealed that, as for normal (native adult-directed) speech, the durational contrast was smaller for short vowels than for long vowels. Also, there was a tendency for the durational contrast to be smaller for short vowels than for long vowels in FDS but not in LS. The overall increase in the duration of vowels is consistent with previous findings (Scarborough et al. 2007; Ashby, 2004) who also found vowels significantly longer in FDS. Biersack et al. (2005) found only a trend for longer vowels in FDS and Knoll et al. (2009) found no difference in vowel duration but this may be due to the fact that in both studies the foreign listeners were imaginary. On the contrary, our study involved real listeners which provided real interaction and feedback to the speaker. Our results for LS are in line with the findings of Junqua (1993) and Ashby (2004) who also found an increase in the duration of vowels.

To our knowledge, previous research in foreigner-directed speech has focused mainly on the durational analysis of vowels (Scarborough et al. 2007; Biersack et al. 2005). Our study extends the findings to the case of plosive consonant duration. Analysis of the plosives showed that durational differences between voiced and voiceless consonants was substantially enhanced for FDS by nearly 50%. In contrast to the results obtained for FDS, the durational contrast in Lombard speech for plosives was reduced. As far as the mean duration of consonants is concerned, we found an increase in LS. These findings extend the results reported by Lu (2010) to the case of non-alveolar plosives. (Lu found an increase of the alveolar plosives only.) Our study shows some variability in acoustic changes from one speaking style to another. The overall increase in duration suggests that foreigner-directed speech is slower than normal and Lombard speech. We also assume that foreigner-directed speech may be more intelligible than both adult-directed and Lombard speech in quiet conditions.

However, since it has been reported that Lombard speech is more intelligible than speech produced in quiet when both are presented in equivalent amounts of noise (Dreher and O'Neill 1957; Summers et al. 1988; Lu and Cooke 2008) we are unable to say how the changes seen in FDS affect intelligibility in noisy conditions. Further studies are needed to investigate potential complementary and antagonistic interactions between FDS, noise-induced speech, and phonological contrasts. Although our study involved a limited number of subjects and the task elicited read speech, the findings support the view that speakers adapt their speech to meet the needs of the target audience. Future studies on intelligibility-enhancing speech modes should involve communicative tasks that elicit more natural speech. Also, since both FDS and Lombard speech are intelligibility-enhancing modes of speech, future perception tests should show which speech styles and which phonetic details are responsible for intelligibility increases in non-native listeners in quiet and noise.

5. Acknowledgements

This work was supported by the Marie Curie RTN, Sound2Sense (S2S). The authors would like to thank Jon Barker and Guy Brown (University of Sheffield, UK) for help with recordings.

REFERENCES

- Ashby, S. 2004. Is infant-directed speech hyperspeech? An acoustical analysis of speech to infants and other accommodative speech styles. (PhD dissertation, University College Dublin.)
- Biersack, S., V. Kempe and L. Knapton. 2005. "Fine-tuning speech registers: a comparison of the prosodic features of child-directed and foreigner-directed speech". *Proceedings of the 9th European Conference on Speech Communication and Technology, Lisbon*. 2401–2405.
- Boersma, P. and D. Weenink. 2005. Praat: Doing phonetics by computer. (Version 4.3.04.) [Computer program.] <<http://www.praat.org>>
- Cruttenden, A. 2001. *Gimson's pronunciation of English*. Oxford: Hodder Arnold.
- DePaulo, B. and L. Coleman. 1986. "Talking to children, foreigners, and retarded adults". *Journal of Personality and Social Psychology* 51(5). 945–959.
- Dreher, J.J. and J.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.

- Garnier, M., L. Bailly, M. Dohen, P. Welby and H. Løevenbruck. 2006. "An acoustic and articulatory study of Lombard speech: Global effects on the utterance". *Proceedings of the Interspeech/ICSLP 2006*, Pittsburgh.
- Gimson, A.C. 1989. *An introduction to the pronunciation of English*. London: Routledge.
- Junqua, J.C. 1993. "The Lombard reflex and its role on human listeners and automatic speech recognizers". *Journal of the Acoustical Society of America* 93. 510–524.
- Knoll, M.A. and M. Uther. 2004. "'Motherese' and 'Chinese': Evidence of acoustic changes in speech directed at infants and foreigners". *Journal of the Acoustical Society of America* 116. 2522.
- Knoll, M., L. Scharrer, and A. Costall. 2009. "Are actresses better simulators than female students? The effects of simulation on prosodic modifications of infant- and foreigner-directed speech". *Speech Communication* 51. 296–305.
- Lombard, E. 1911. "Le signe de l'élévation de la voix" [The sign of the rise in the voice]. *Ann. Maladies Oreille, Larynx, Nez, Pharynx* 37. 101–119.
- Lu, Y. and M.P. Cooke. 2008. "Speech production modifications produced by competing talkers, babble and stationary noise". *Journal of the Acoustical Society of America* 124. 3261–3275.
- Lu, Y. 2010. Production and perceptual analysis of speech produced in noise. (PhD dissertation, University of Sheffield.)
- Patel, R. and K.W. Schell. 2008. "The influence of linguistic content on the Lombard effect". *Journal of Speech, Language, and Hearing Research* 51. 209–220.
- Sankowska, J., M. Cooke and M. García Lecumberri. 2009. "Interaction of intrinsic vowel and consonant durational correlates with Lombard speech". Paper presented at the 4th Phonetics and Phonology in Iberia (PaPI) Conference, Las Palmas de Gran Canaria.
- Scarborough, R., J. Brenier, Y. Zhao, L. Hall-Lew, and O. Dmitrieva. 2007. "An acoustic study of real and imagined foreigner-directed speech". *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken. 2165–2168.
<[http:// www.icphs2007.de](http://www.icphs2007.de)>
- Summers, W., D.B. Pisoni, R.H. Bernacki, R.I. Pedlow and M.A. Stokes. 1988. "Effects of noise on speech production: Acoustic and perceptual analysis". *Journal of the Acoustical Society of America* 84. 486–490.
- Uther, M., M.A. Knoll and D. Burnham. 2007. "Do you speak E-n-g-l-i-s-h? A comparison of foreigner- and infant-directed speech". *Speech Communication* 49. 2–7.
- Wells, J.C. 1981. *Accents of English*. Cambridge: Cambridge University Press.

Corresponding author:

Joanna Sankowska
Facultad de Letras
Universidad del País Vasco
Paseo de la Universidad 5
Apdo 2111
01006 Vitoria
Spain
j.sankowska@hotmail.com