AN ARTICULATORY PROSODY STUDY OF /u/: MOTOR EQUIVALENCE

Marija Tabain [(a)], Pascal Perrier [(b)], Christophe Savariaux [(b)] and Richard Beare [(c)]

(a) MARCS Auditory Laboratories, University of Western Sydney, AUSTRALIA.
(b) Institut de la Communication Parlée, Grenoble, FRANCE.
(c) Mathematical and Information Sciences, CSIRO, Sydney, AUSTRALIA.

m.tabain@uws.edu.au

Abstract

Acoustic and articulatory data are presented for the vowel /u/ at 5 different prosodic boundaries in French. It is shown that although the 3 speakers studied produce similar acoustic output - with lower F1 and F2 at stronger prosodic boundaries – they adopt different articulatory strategies to achieve these acoustic targets. Two of the speakers use a combination of tongue backing/raising and lip protrusion, whereas the third speaker primarily uses the tongue tip. It is suggested that these strategies represent motor equivalence in the production of /u/, and it is hoped that the data presented here will lead to better articulatory models of this vowel.

1. Introduction

Many studies have shown an effect of prosodic structure on the articulation and acoustics of individual consonants and vowels [Fougeron & Keating 1997; Byrd & Saltzman 1998; Byrd 2000; Cho & Keating 2001; Fougeron 2001; Cho 2002; Keating et al. 2003; Tabain 2003a, 2003b; Tabain & Perrier, in press]. Broadly speaking, duration is greater, and articulation and formant structure are more prototypical, at stronger prosodic boundaries (e.g. Utterance or Intonational phrase); and duration is shorter, with more reduced articulation and centralized vowel formants, at weaker prosodic boundaries (e.g. Word or Syllable). For example, there is greater linguo-palatal contact for /n/ at a stronger prosodic boundary than at a weaker prosodic boundary (Fougeron & Keating, 1997), while /a/ is produced with a lower tongue and jaw position at stronger prosodic boundaries than at weaker prosodic boundaries (Tabain, 2003b). This effect of prosodic structure is, however, not identical across segments – for instance, segments which are intrinsically resistant to variability and coarticulation, such as /i/ and /s/, show less effect from prosodic structure than do other segments (Fougeron, 2001; Cho, 2002; Tabain & Perrier, in press)

In this study, we examine the acoustics and articulation of the vowel /u/ in French, since to our knowledge, the behaviour of /u/ - one of the 3 "point" vowels - at different prosodic boundaries has not been investigated. This vowel involves a high tongue position to the rear of the oral cavity, in contrast to /i/ which has a more front (high) tongue position, and to /a/ which has a low tongue position. /u/ also involves significant rounding of the lips, in contrast to both /i/ and /a/. The gestures involved in the production of /u/ are therefore quite different to those required for the other point vowels which we have studied within a prosodic framework.

2. Method

Three <u>speakers</u> of metropolitan French were recorded at the ICP studios in Grenoble. Two of the speakers (AV [female] and CV [male]) had participated in previous studies on the articulatory prosody of the vowel /a/ (Tabain, 2003a & 2003b) and of the vowel /i/ (Tabain & Perrier, in press). Speaker LN (female) was new to this series of studies.

<u>Acoustic and articulatory data</u> were recorded simultaneously. Acoustic data were sampled at 20 kHz; and articulatory data were sampled at 500 Hz using the 10-channel Carstens EMA system. Four EMA sensors were placed on the tongue (from back to tip); one each on the upper and lower lips; one on the jaw (i.e. lower gum); and two were used as reference points (upper gum and nose). Rotation of the articulatory data to the occlusal plane, and subtraction of the reference points, were carried out in MATLAB. Acoustic data were labelled using the EMU speech labeller (Cassidy & Harrington, 2001). All other analyses, including smoothing of the articulatory data using the regression-based Lowess filter, were carried out using the R statistical package (R Development Core Team, 2003).

<u>Stimuli</u> consisted of the following sentences (sentences are listed in order of prosodic strength, where

Proceedings of the 10th Australian International Conference on Speech Science & Technology

Accepted after full paper review

Utterance is the strongest boundary and Syllable is the weakest boundary – cf. Nespor & Vogel, 1986):

1. Paul aime Papou. **B**ouba les protège en secret.
Utterance
Paul loves Papou. Bouba looks after them in secret
2. Le pauv' Papou, **B**ouba et Paul arriveront demain.
Intonational
Poor Papou, Bouba and Paul are coming tomorrow
3. Tonton, Papou, **B**ouba et Paul arriveront demain.
Accentual
Uncle, Papou, Bouba and Paul are coming tomorrow
4. Paul et Papou **B**ouba arriveront demain.
Word
Paul and Papou Bouba are coming tomorrow
5a. Tonton et Papou**b**ou arriveront demain.
Syllable
Uncle and Papoubou are coming tomorrow
[Speaker AV]
5b. Les belles Papou**b**oubas arriveront demain.
Syllable
The beautiful Papouboubas are coming tomorrow
[Speakers CV & LN]

Note that the target vowel /u/ is placed in the sequence /apu # Cu/. The consonant (bold type in the sentences listed above) was varied to be one of /b d g f s S/. Each speaker produced 10 repetitions of the stimuli set, resulting in about 300 /u/ tokens per speaker (5 prosodic boundary contexts X 6 consonant contexts X 10 repetitions). Note that the Syllable sentence context differs for the 3 speaker speakers, due to planning error (the nuclear accent of the sentence is further away from the target syllable in sentence 5b than in sentence 5a; also, the target syllable is the 5th syllable in the sentence in sentence 5a, rather than the 4th syllable, as in all the other sentences). Note also that /u/ in the Utterance boundary context is followed by a pause, whereas in all other contexts it is followed by a consonant.

Speakers tended to produce the Utterance boundary with a pause, and the Intonational boundary without a pause but with a continuation contour. The Accentual phrase boundary tended to be produced as a list.

For all of the results presented below, the start of the vowel /u/ was taken at the acoustic release of the preceding consonant /p/, and the end of the vowel was taken at the acoustic endpoint of the vowel. Results will be collapsed across consonant contexts.

3. Results

Figure 1 presents the results for acoustic vowel duration of /u/ at different prosodic boundaries. It can be seen that, consistently with previous studies, duration is greater at stronger boundaries and shorter at weaker boundaries. This is true for all 3 speakers, although speakers differ in which prosodic boundaries are grouped together.

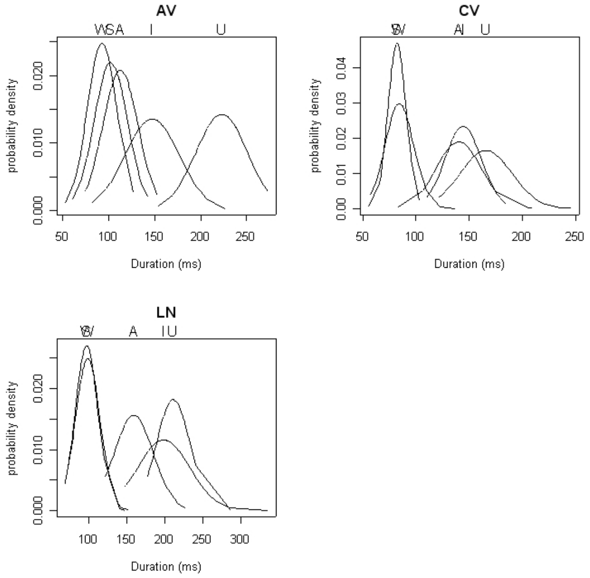


Figure 1: Normal distribution plots of /u/ acoustic vowel duration at different prosodic boundaries. Data for 3 speakers of metropolitan French. For this and all following figures, "U" = Utterance, "I" = Intonational phrase, "A" = Accentual phrase, "W" = Word and "S" = Syllable.

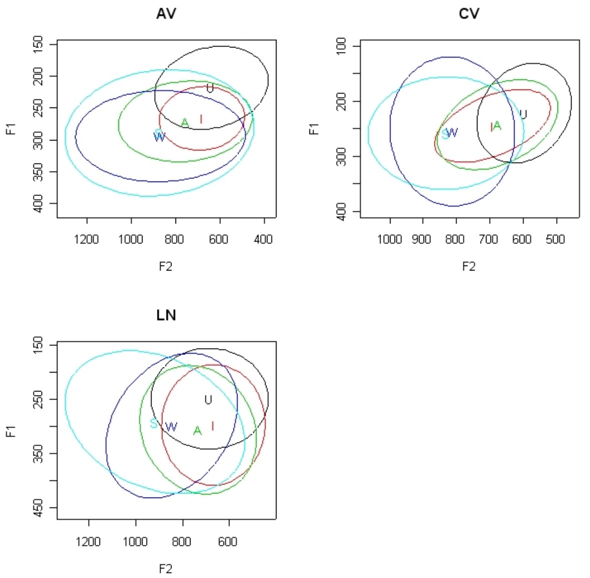


Figure 2: Ellipse plots showing means and 2.45 standard deviations for F1 and F2 (in Hz) of the vowel /u/ at different prosodic boundaries.

Figure 2 presents ellipse plots of the acoustic vowel space for /u/ at different prosodic boundaries. It can be seen that for all 3 speakers, there is a tendency for both F1 and F2 to be lower at stronger prosodic boundaries, and higher at weaker prosodic boundaries.

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004. Copyright, Australian Speech Science & Technology Association Inc.

Accepted after full paper review

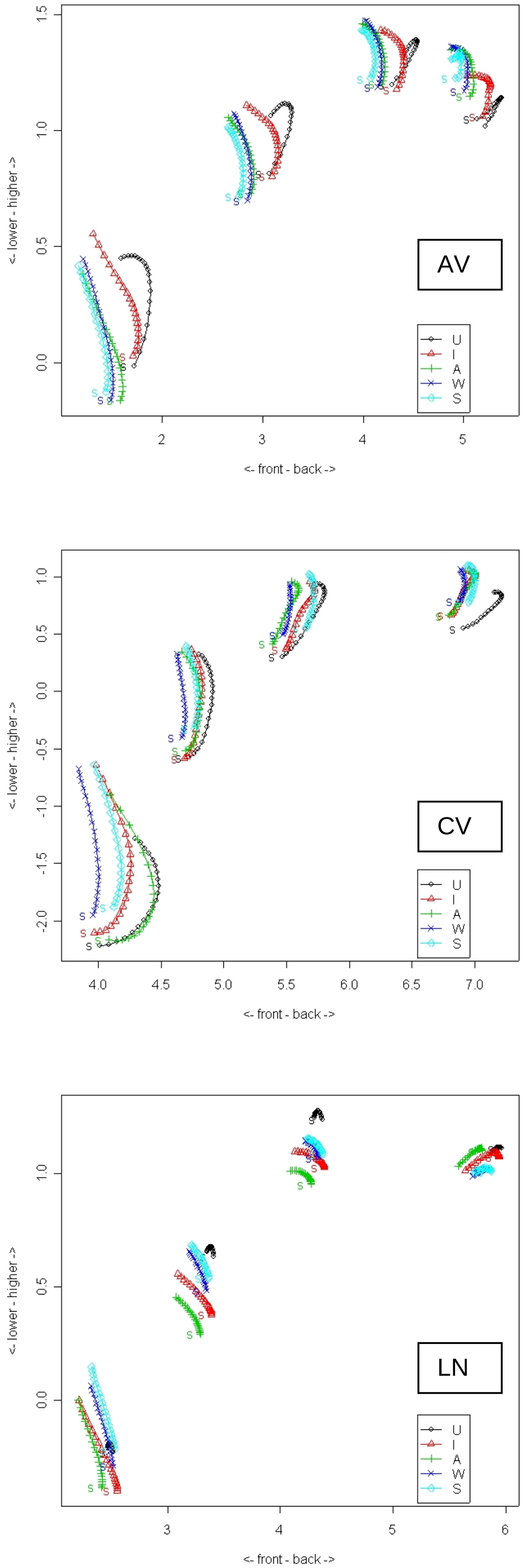
It may be concluded at this point that a low F1 and a low F2 are the acoustic goal for a prototypical /u/ (as evidenced by the Utterance boundary data in Figure 2, and as noted in many textbooks on speech – e.g. Johnson, 1997). We assume here that F1 is the Helmholtz resonance formed by the constriction between the tongue and the palate, and the pharyngeal cavity; and that F2 is the Helmholtz resonance formed by the constriction at the protruded lips, and the oral cavity. Thus, in order to lower the Helmholtz resonance of either constriction-cavity, a speaker may (a) *decrease* the area of the constriction; (b) *increase* the area of the cavity; (c) *increase* the length of the constriction; and/or (d) *increase* the length of the cavity.

Figure 3 presents articulatory data from the 4 tongue transducers. Note that the shape of the Utterance boundary data differs to the shape of the other boundary data, since the Utterance boundary is followed by a pause (usually a "neutral" tongue position), whereas the other boundaries are followed by a consonant. Note also that since data in Figure 3 are time-normalized, the mid-point in each trajectory (there are 20 points in each trajectory, so that the mid-point is between the 10^{th} and 11^{th} point) corresponds in time to the acoustic vowel data shown in Figure 2 (also taken at the temporal mid-point). Data for the lips are not shown due to space considerations, although they will be briefly described in the discussion which follows.

It can be seen that speaker AV has a higher and more back tongue body at stronger prosodic boundaries; although not shown here, her lips are also more protruded at stronger prosodic boundaries. Such strategies serve to decrease the area of the constriction and hence lower both F1 and F2.

Speaker CV, by contrast, although showing some tongue backing at the stronger boundaries, primarily uses the tongue tip to signal stronger prosodic boundaries. This serves to increase the size of the resonating cavity and hence lower the resonance frequency. There were minimal effects of prosodic boundary on lip protrusion for this speaker.

Finally, speaker LN shows a much higher tongue body for the stronger prosodic boundaries; however, whereas for the other speakers, the tongue became progressively lower (and/or more forward) as the prosodic boundary became weaker, this is only true as far as the Intonational and Accentual phrase boundaries for this speaker. For the Word and Syllable boundaries, the tongue returns to a higher position. A similarly unusual pattern was observed for the tongue data for this speaker. We are currently unable to fully explain this result for this speaker.



Accepted after full paper review

Figure 3 (previous page): Plots showing time-normalized and averaged tongue movement trajectories for the vowel /u/ at different prosodic boundaries. Data are from 4 EMA sensors. "S" marks the start of the tongue movement, measured from the acoustic release of the preceding consonant (taken to the acoustic end of the vowel). The back of the tongue is to the top right of each figure, and the tongue tip is to the bottom left of each figure.

4. Discussion

Although similar acoustic patterns were observed for all 3 speakers in this study, each speaker adopted a different articulatory strategy to achieve this acoustic goal. These results recall results from motor equivalence studies on /u/ (e.g. Perkell, Matthies, Svirsky & Jordan, 1993; Savariaux, Perrier & Orliaguet, 1995; de Jong, 1997; Savariaux, Perrier, Orliaguet & Schwartz, 1999). However, in contrast to these previous studies, the data from this study were generated neither from rate manipulation of speech, nor from bite block experiments: instead, linguistic theory was used to generate sentences where the duration of the vowel varied naturally, hence controlling the likelihood of achieving the prototypical vowel target. We thus hope to have provided extra insight into production of the vowel /u/, which has been relatively less studied that the other point vowels.

5. Acknowledgements

We would like to thank our speakers for participating in this study. This research was supported by an Australian Research Council fellowship to the first author.

6. References

Byrd, D. (2000). *Articulatory vowel lengthening and coordination at phrasal junctures*. Phonetica, 57: 3-16.

Byrd, D. & Saltzman, E. (1998). *Intragestural dynamics of multiple prosodic boundaries*. Journal of Phonetics, 26: 173-199.

Cassidy, S. & Harrington, J. (2001). *Multi-level annotation in the EMU speech database management system.* Speech Communication, 33: 61-77.

Cho, T. (2002). *Effects of Prosody on Articulation in English.* New York, Routledge.

Cho, T. & Keating, P. (2001). *Articulatory and acoustic studies on domain-initial strengthening in Korean*. Journal of Phonetics 29: 155-190.

De Jong, K. (1997). *Labiovelar compensation in back vowels.* Journal of the Acoustical Society of America, 101: 2221 - 2233.

Fougeron, C. (2001). *Articulatory properties of initial segments in several prosodic constituents in French*. Journal of Phonetics, 29: 109-135.

Fougeron, C. & Keating, P. (1997). *Articulatory strengthening at edges of prosodic domains*. Journal of the Acoustical Society of America, 101: 3728-3740.

Johnson, K. (1997). *Acoustic and auditory phonetics.* Oxford, Blackwell.

Keating, P., Cho, T., Fougeron, C. & Hsu, C-S (2003). *Domain-initial articulatory strengthening in four languages*. In J. Local, R. Ogden, & R. Temple (eds), Papers in Laboratory Phonology VI (pp. 143-161). Cambridge, U.K.: Cambridge University Press.

Nespor, M. & Vogel, I. (1986). *Prosodic phonology*. Dordrecht, Foris.

Perkell, J., Matthies, M., Svirsky, M. & Jordan, M. (1993). *Trading relations between tongue-body raising and lip-rounding in production of the vowel /u/: a pilot "motor equivalence" study*. Journal of the Acoustical Society of America, 93: 2948-2961.

R Development Core Team (2003). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, URL http://www.R-project.org.

Savariaux, C., Perrier, P. & Orliaguet, J-P. (1995). *Compensation strategies for the perturbation of the rounded vowel [u] using a lip-tube: A study of the control space in speech production*. Journal of the Acoustical Society of America, 98: 2428–2442.

Savariaux, C., Perrier, P., Orliaguet, J-P. & Schwartz, J-L. (1999). *Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis*. JASA, 106: 381-393.

Tabain, M. (2003a). *Effects of prosodic boundary on /aC/ sequences: acoustic results*. Journal of the Acoustical Society of America, 113: 516-531.

Tabain, M. (2003b). *Effects of prosodic boundary on /aC/ sequences: articulatory results*. Journal of the Acoustical Society of America, 113: 2834-2849.

Tabain, M. & Perrier, P. (in press). *Articulation and acoustics in pre-boundary position in French*. Journal of Phonetics

Accepted after full paper review