Articulatory Characteristics of Function Words in English: a Case Study

Mitsuhiro Nakamura

Nihon University, Tokyo, Japan
E-mail: nakamura.mitsuhiro99@nihon-u.ac.jp

Abstract

This study reports on the continuing investigation into the articulatory realisation of function words in English. Using data from the Multichannel Articulatory (MOCHA) database, the consonant articulations in function and content words were examined as a function of position in an utterance. The results of the EMA data analysis for [t] and [n] showed that the tongue tip movement in the two word classes differed temporally, not spatially. This word class distinction is considered from the viewpoint of the interface between prosody and articulation.

1 Introduction

Position in the prosodic structure is regarded as a linguistic factor generating variation in the articulatory properties of phonetic segments [1]. As is well known, English function words exhibit patterns of phonetic variation depending on utterance position. Monosyllabic function words, as opposed to content words, are commonly realised as a stressless weak form, but, when isolated or at the beginning or end of a sentence, they are realised as a stressed strong form. The distinction between these two phonetic forms has been shown to correlate with the difference in vowel quality, namely full vs. reduced or schwa. Little, however, is known about the mechanism involved in the consonant articulation of these two forms.

This research is an attempt to uncover whether and how the articulatory properties of the consonant vary between the strong and weak forms and between the word classes, content vs. function. This will allow us to assess the feasibility of integrating the word class distinction into the interface between prosody and articulation.

2 Methodology

The speech materials used in this study were drawn from the Multichannel Articulatory (MOCHA) database [2]: Carstens Electromagnetic Articulograph (EMA, 500Hz sample rate), Laryngograph, and Reading Electropalatograph were used for articulatory data acquisition. The consonants examined were [t] and [n]. The tongue tip movement was analysed for the selected utterances spoken by three speakers of Southern British English (RP) denoted below as SE, SA, and AP.

Two sets of data were prepared. Dataset 1 consisted of a function word 'to' and monosyllabic content words with the onset consonant [t] (e.g. *take, two,* and *teach*). Dataset 2 consisted of the word 'in': this function word has no separate weak form in RP [3].

For both sampling contexts, three positions in an utterance were identified: (absolute) initial, medial, and (absolute) final. When sampling in the medial position, phonetic contexts were specified. For dataset 1, the target item was preceded by a vowel (e.g. 'Try *to*,' 'ability *to*,' 'who *teach*,' and 'I *took*'). For dataset 2, the item was followed by a vowel (e.g. '*in* a lively' and '*in* athletic'). The MOCHA corpus was searched for examples which met the above requirements, giving 26 tokens per speaker for dataset 1 and 9 tokens per speaker for dataset 2. A few samples were eliminated from the analysis due to immeasurability of the EMA tongue tip data. The number of tokens for each subject in the two datasets is summarised in Tables 1 and 2 (no token was found for 'to' in the final position in dataset 1).

Table 1. *Number of items in dataset 1*

Speaker	Content words			'to'		Total
	Initial	Medial	Final	Initial	Medial	
SE	5	7	2	1	11	26
SA	5	7	2	1	11	26
AP	4	7	2	1	11	25

Table 2. *Number of items in dataset 2*

Speaker	Initial	Medial	Final	Total
SE	1	6	2	9
SA	1	5	2	8
AP	1	3	1	5

"The best way to learn is to solve extra problems."
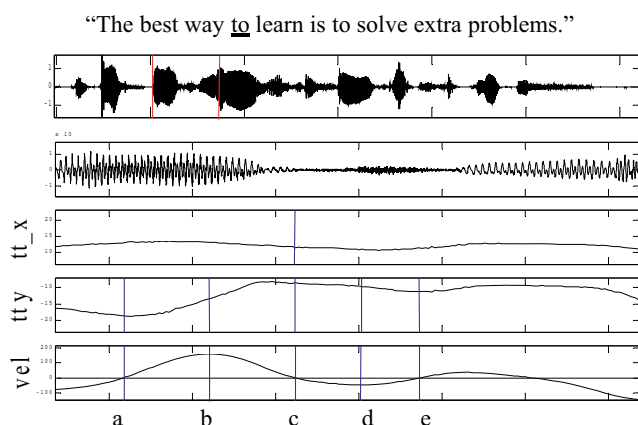
Figure 1: *Measurement points*

Table 3. *Speaker and word class effects on TTy*

	Initial position		Medial position	
Speaker	F(2,13)	34.93**	F(2,50)	102.68**
Word class	F(1,13)	3.07	F(1,50)	4.53*

unmarked=non–significant; *p<0.05; **p<0.01

Table 4. *Speaker and position effects on TTy*

	Content words		Function word 'to'	
Speaker	F(2,36)	101.27**	F(2,32)	82.81**
Position	F(2,36)	2.02	F(1,32)	0.18

unmarked=non–significant; **p<0.01

The analysis of the EMA TT coil (placed at about 7-10mm posterior to the tip) was performed using the EMA tools [2]. Figure 1 shows sample data with the measurement points marked. Following the criteria used by [4], time and position for consonant onset (point a), target (point c), and end (point e) were defined by zero-crossings of the velocity trajectory. Also, the time of peak velocity for both the closing (point b) and opening (point d) movements were recorded. Based on these points and positions, the following spatiotemporal variables were calculated. *Constriction peak* was the peak displacement of the tongue tip's vertical movement (TTy) at target time point (c): this is specified as the distance from the upper-incisor reference coil. *Total duration* was measured for the temporal interval between the time point (a) and (e). *Closing* and *opening* durations were the intervals (a)-(c) and (c)-(e) respectively. Statistical comparison was made by one- and two-way ANOVAs with Spjotvoll/Stoline post-hoc test (p<0.05); Sheffé's test (p<0.05) was used for the analysis of the position effects on the function word 'to' in dataset 1.

3 Analysis of dataset 1

In this analysis we compare the articulatory movement for [t] in 'to' with that in the onset of monosyllabic content words as a function of utterance position and speaker. Figure 2 presents the mean x/y coordinates of peak displacement for the three speakers.

The difference between the word classes was examined for the initial and medial positions. The results in Table 3 reveal that the TTy locations vary significantly between the speakers. For both positions, SA has a higher TTy location: SA>SE=AP. A significant difference was also found in the medial position. However, the mean values are not substantially divergent between the content word tokens (−5.3mm) and the function word tokens (−5.9

mm): this turned out to be barely non-significant (p=0.059).

The difference between the utterance positions was tested separately for each word class. Table 4 reveals that the TTy locations do not vary between the utterance positions but between the speakers: for the content words, SA>AP>SE; for the function word, SA>SE=AP.

Overall, speakers vary idiosyncratically in their degree of the constriction peak: SA has a markedly higher TTy position. In contrast, the distinction between the word classes and between the utterance positions is not reflected clearly and significantly.
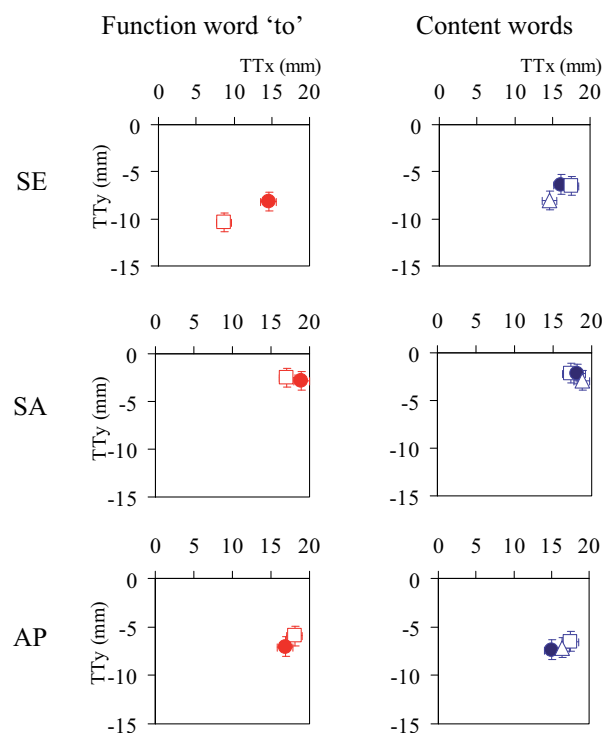

Figure 2: *Mean TTx/y displacement for [t]*

□=initial, ●=medial; □=initial, ●=medial, △=final
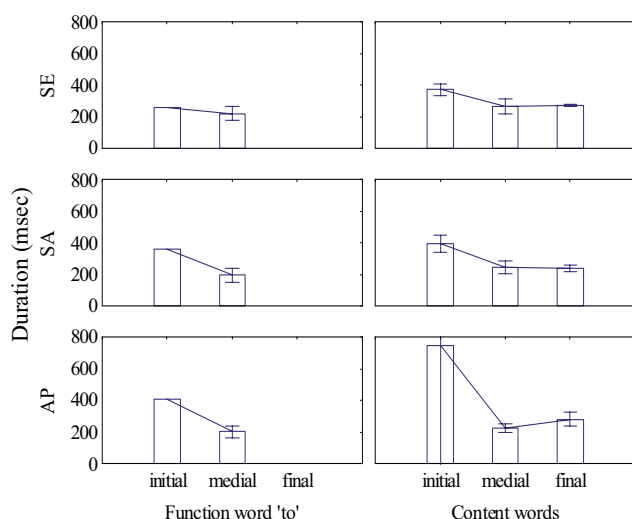Error bars indicate ±1 standard deviation.

Figure 3: *Mean total duration*
Error bars indicate ±1 standard deviation.

Figure 4: *Mean closing and opening duration*
Error bars indicate ±1 standard deviation.

We now turn to the mean total duration in Figure 3 and the mean closing and opening duration in Figure 4. The pattern of the three durations is similar across the speakers, but it varies with the word class. Statistical results are shown in Table 5 below.

Table 5. *Speaker and word class effects on three durations of [t]*

	df	Total	Closing	Opening
Speaker	F(2,73)	0.47	0.10	0.45
Word class	F(1,73)	6.00*	11.27**	3.74*

unmarked=non-significant; *p<0.05; **p<0.01

Table 6 presents the results of the effect of word class on the three durations. There is no significant difference between the word classes in the initial position. In the medial position, however, the total and opening duration of the content word tokens was longer than that of the function word. It is supposed that the longer opening duration is related to the stronger aspiration in the medial content words.

Table 7 indicates the results of the position effects.

Table 6. *Word class effects on three durations*

(a) Initial position

	Total	Closing	Opening
F(1,15)	0.32	2.35	0.18
Relation	C = F	C = F	C = F

(b) Medial position

	Total	Closing	Opening
F(1,52)	10.55**	2.13	8.62**
Relation	C > F	C = F	C > F

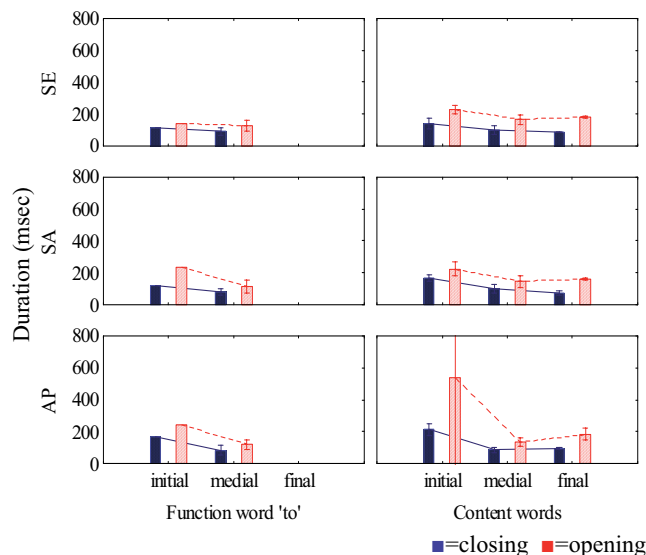unmarked=non-significant; **p<0.01
C=Content words, F=Function word 'to'

Both content and function words show that all the durations are significantly longer in the initial position. One exception is the opening duration of the content word: the difference levels out across the three positions. This appears to serve as the basis for the invariable realisation of aspiration in the onset [t] regardless of the utterance positions. In contrast, the opening duration and the other two durations are shortened in the medial function word.

4 Analysis of dataset 2

In this analysis, we examine the tongue tip movement for the coda [n] in the word 'in' as a function of utterance position and speaker. Figures 5 and 6 reveal the mean x/y coordinates of peak displacement and the mean total, closing, and opening duration for the three speakers respectively.

Similar to the results of the analysis of dataset 1, the TTy location differs significantly between the speakers [F(2,17)=50.06, p<0.01] but not between the positions [F(2,17)=1.33, p=0.28]. In contrast, as

Table 7. *Position effects on three durations*

(a) Content words

	Total	Closing	Opening
F(2,38)	4.12*	31.83**	1.97
Relation	I > M=F	I > M=F	I = M = F

(b) Function word 'to'

	Total	Closing	Opening
F(1,34)	24.68**	10.33**	14.49**
Relation	I > M	I > M	I > M

unmarked=non-significant; *p<0.05; **p<0.01
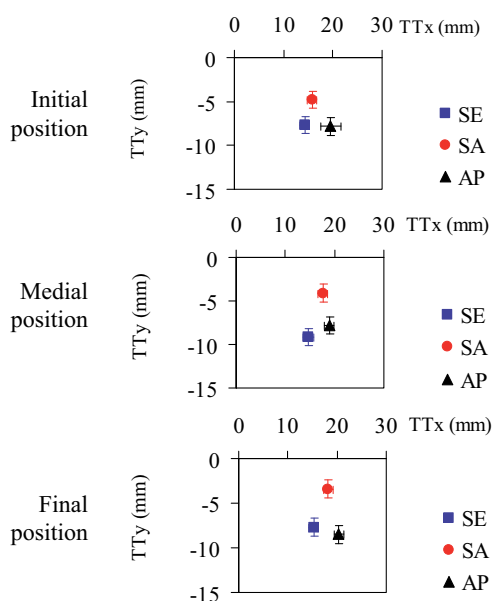I=initial, M=medial, and F=final position

Figure 5: *Mean TTx/y displacement for [n]*
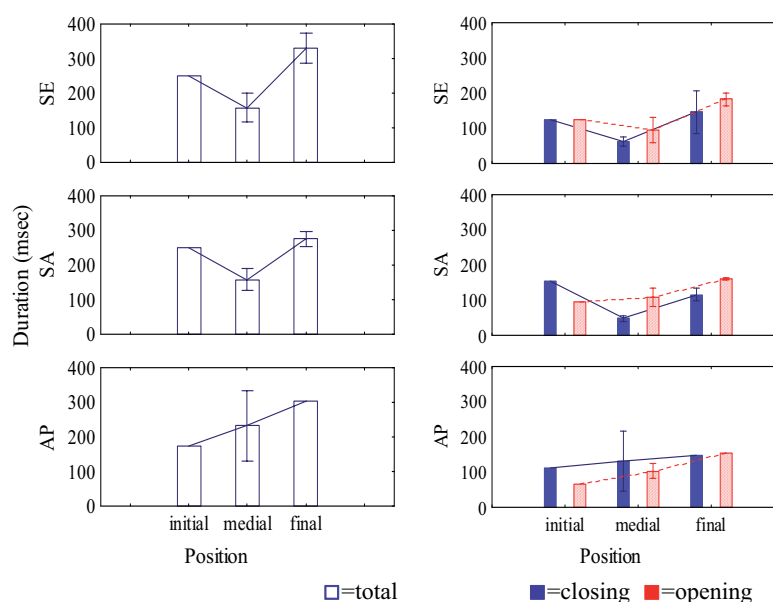Error bars indicate ±1 standard deviation.

Figure 6: *Mean total duration (left) and closing and opening duration (right) for [n]*
Error bars indicate ±1 standard deviation.

is shown in Table 8, the speaker effect is non-significant in the results of the analysis of the three durations. Two patterns are actually observed: increase in the initial and final positions; and progressive increase from the initial to the final position. Statistically, the closing duration is found to be shorter in the medial position. Also, the total and opening durations are longer in the final position, a pattern which reflects the effect of final lengthening of the coda [n].

5 Discussion and conclusion

The results presented above are compatible with previous studies on positionally-conditioned variations in the consonant articulation (e.g. [1, 4, 5]). Both the word class distinction and the positional variation can be sought in the temporal domain. Spatial variations in terms of peak displacement are largely idiosyncratic. For both function and content words, temporal lengthening in the initial position and shortening in the medial position were consistently found. This lengthening, or shortening, serves as a basis for the alternation between the strong and weak forms of function words. However,

it was found that the opening duration of [t] in the content word was not affected by utterance position. Therefore, prosodic position effects are realised similarly on the one hand but differently on the other, suggesting that differential articulatory encoding of the two word classes is involved in speech production. Further research is needed to understand the role of the content/function word distinction in the interface between prosody and articulation.

Acknowledgements

References

[1] P. Keating, *et al*. Domain-initial strengthening in four languages. In J. Local, R. Ogden, and R. Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI*. Cambridge University Press, Cambridge: 143-161, 2003.

[2] A. Wrench. The MOCHA-TIMIT Articulatory Database. at http://www.cstr.ed.ac.uk/research/projects/artic/mocha.html, 2000.

[3] J. Wells. *Longman Pronunciation Dictionary*. Third Edition. Pearson Education, Essex, 2008.

[4] D. Byrd, *et al*. Interacting effects of syllable and phrase position on consonant articulation. *J. Acoust. Soc. Am.*, 118(6): 3860-3873, December 2005.

[5] D. Byrd, *et al*. Phrasal signatures in articulation. In M. Broe and J. Pierrehumbert (eds.), *Acquisition and the Lexicon: Papers in Laboratory Phonology V*, Cambridge University Press, Cambridge: 70-88, 2000.

Table 8. *Speaker and position effects on three durations of [n]*

	df	Total	Closing	Opening
Speaker	F(2,17)	0.68	2.50	0.40
Position	F(2,17)	9.84**	5.36*	10.76**
	Relation	F > I=M	I=F > M	F > I=M

unmarked=non–significant; *p<0.05; **p<0.01
I=initial, M=medial, and F=final position