PROMINENCE ENHANCES VOICELESS-NESS AND NOT PLACE DISTINCTION IN ENGLISH VOICELESS SIBILANTS

Meghan Clayards[a] *& Thea Knowles*[b]

[a]McGill University, Canada; [b]University of Western Ontario, Canada
meghan.clayards@mcgill.ca; tknowle3@uwo.ca

ABSTRACT

Prominence and prosodic strengthening make (at least some) segmental articulations more extreme, potentially enhancing acoustic contrasts. Here we found that prominence alone does not modulate the acoustic distinctiveness of the English sibilant place (/s/ vs. /ʃ/) contrast, although speakers can produce the sibilants more distinctively when contrasting with a minimal pair. Furthermore, we found evidence that sibilants at the end of prominent syllables were produced with an earlier glottal opening gesture. This produced more breathiness in the vowel and often an [h]-like segment between the vowel and the sibilant. This is consistent with the observation that prominence increases the glottal opening gestures as has been previously observed for prosodic strengthening of initial voiceless stops. Here we extend this observation to sibilants as well as codas of prominent syllables. Our results also suggest that prominence modulates gestural constrictions and not acoustic contrasts directly.

Keywords: prominence, segmental distinctiveness, voiceless sibilants, voice quality.

1. INTRODUCTION

Segmental articulations have been argued to be lengthened and strengthened in prosodically strong positions, such as phrase-initially [14, 8] and under emphasis [11, 17, 9], potentially enhancing the acoustic cues signaling segmental contrasts [1]. For example, word initial stop voicing contrasts are enhanced under prominence [11, 17, 9] and vowels are less centralized [7, 9]. However, prosodic strengthening does not necessarily enhance all acoustic contrasts – Cole et al. [11] found that stop voicing but not place was enhanced under prominence – nor is the articulation of all segments equally sensitive to prosodic modulation. The amount of lingual contact for /s/ has been shown to vary less according to prosodic position than other segments in French [13] and English [19] possibly because of the rigid articulatory requirements of sibilant production especially with respect to constriction degree [27]. Thus the effect of prominence and prosodically conditioned variation on segmental contrasts is not straightforward.

One possibility is that prominence does not enhance contrasts directly, but rather affects the degree of articulatory constriction including glottal opening [18] as well as having a general lengthening effect [14]. This would indirectly enhance voicing contrasts but not place contrasts because they rely on constriction location. More generally prominence should only enhance contrasts that rely on constriction degree (including voicing) similar to traditional notions of fortition. Clear speech, on the other hand, may more directly modulate phonological contrasts [24, 17, 5] (though see [16, 26] for evidence that clear speech doesn't always enhance contrasts). In this paper we test two hypotheses:

(1) Prominence modulates acoustics through strengthening gestures (especially glottal opening) not by enhancing contrasts directly

(2) Careful speech enhances acoustic contrasts

Examination of voiceless stops doesn't allow us to distinguish hypothesis (1) from a direct contrast enhancing role of prominence as both would expand the difference between voiced and voiceless stops. Voiceless sibilants, however, provide both a lingual constriction and a glottal opening as possible targets of enhancement. The contrast between /s/ and /ʃ/ is not one of lingual constriction degree, however, but of positioning. Thus what is normally thought to be the primary cue to the place contrast (Centre of Gravity COG of the sibilant noise spectrum) may be resistant to prosodic modulation alone (though see [6, 10] for enhancement of a sibilant place contrast in Mandarin with prominence, at least for some talkers). Furthermore, voiceless sibilants include a glottal opening gesture substantial enough to introduce breathiness in the preceding vowel [22, 20] and in French, vowels preceding /ʃ/ were found to be breathier than those preceding /s/ [25] indicating a possible role for glottal opening in signaling the place contrast. Thus voiceless sibilants represent an interesting test case for the role of prominence versus clear speech in the realization of

acoustic contrasts. Hypothesis (1) predicts that the breathiness of the preceding vowel will be modulated under prominence rather than the COG. Hypothesis (2) predicts that under careful speech the primary cue to the place contrast. i.e. COG will also be modulated.

In this paper we present two experiments which examine the acoustic consequences of prominence and clear speech on the production of voiceless sibilants in English in the onsets and offsets of monosyllabic words. Experiment 1 manipulates prominence alone while Experiment 2 is designed to elicit more clear speech. Our results provide support for both hypotheses.

2. METHODS

2.1. Experiment 1

Participants were asked to read pairs of sentences "as if giving instructions in a game" that consisted of a context sentence and a target sentence including a target word shown in bold in Table 1. The target word was either repeated from the context – where we expected low prominence (*Repeat*), or was under contrastive focus (*Prominent*) – where we expect high prominence (marked by caps), or a filler condition (not discussed here for reasons of space). The sentence containing the target word was always identical across conditions and the target word was always preceded by *the* and followed by *above.*

2.1.1. Design

In Experiment 1a, 21 (13f) native speakers of North American English read sentences containing target C(C)VC words with sibilants in the offset. 36 items were created out of 18 pairs matched for vowel quality (see Table 2 after Discussion section). Each participant read one item in one of the three conditions (*Repeat, Prominent* or filler). In Experiment 1b, 12 participants read sentences containing sibilants in the onsets of CVC words. 18 items were created out of 9 minimal pairs and each participant read all items in all conditions.

Table 1: Example sentence pairs read in the two experiments. The target sentence was always the same within an item. Words in bold are the target words. Words in caps are predicted to be prominent. Target words contained sibilants in either the onset or offset of the word.

Condition	Context	Target
Repeat	Move the leash above the caterpillar	Now, move the **leash** above the orange
Filler	Move the button above the caterpillar	Now, move the **leash** above the orange
Prominent Exp 1	Move the button above the orange	Now, move the **LEASH** above the orange
Prominent Exp 2	Move the lease above orange	Now, move the **LEASH** above the orange

2.1.2. Acoustic Analysis

Word and segment boundaries were force aligned using the prosody lab aligner [15] and sibilant segment boundaries were hand checked and adjusted. All measurements were performed in Praat [4]. To quantify prominence we measured the duration, average intensity and average f0 of each of the target words. Duration and average intensity of the segments were also measured. Spectral centre of gravity (COG) was measured in a 15 ms window at the center of the sibilant noise portion.

For Experiment 1a, we also measured voice quality characteristics of the preceding vowel (H1-H2) and the duration of any voiceless vowel portion for all female speakers. Figure 1 shows an example of a voiceless vowel portion (marked as [h]) observed in our data. Such voiceless portions were also noted in [25] before French sibilants. Harmonics were measured 20 ms before the end of the vowel or 20 ms before the beginning of any voiceless portion.

2.1. Experiment 2

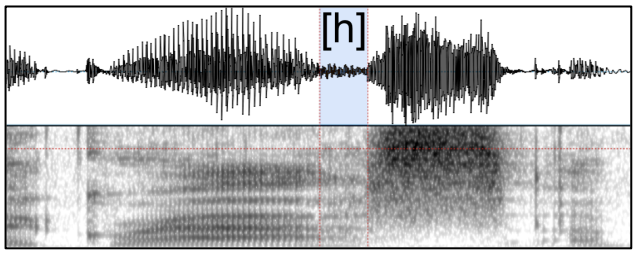
The methods were identical to Experiment 1 except that in the *Prominence* condition the target word also contrasted with a minimal pair in the previous sentence (Table 1). As before, only the target words (bold) were analyzed. 14 (8f) participants read sentences from all three conditions for 36 items. Half of the items had sibilants in the onset and half in the offset.

3. RESULTS

Only the *Repeat* and *Prominent* conditions were analysed. We used mixed effect linear regression for each of the acoustic measurements using the lmer() function in R [2]. For prominence measures, models designed to test whether prominence varied by condition included fixed effects of condition and experiment and their interaction, and random intercepts and slopes for participants and items. For COG and H1-H2, the fixed effects were segment,

condition, experiment and their interactions. A random intercept and a random slope for condition was included for items. A random intercept and slopes for segment, condition and their interaction were included for participants. p values are derived from a interpreting the t-statistic as a z distribution. All fixed effects were centred.

Figure 1: An example production of the word 'grass' showing a section before the /s/ marked as [h] which is voiceless but contains formant structure indicating both the vocal tract and glottis are open.



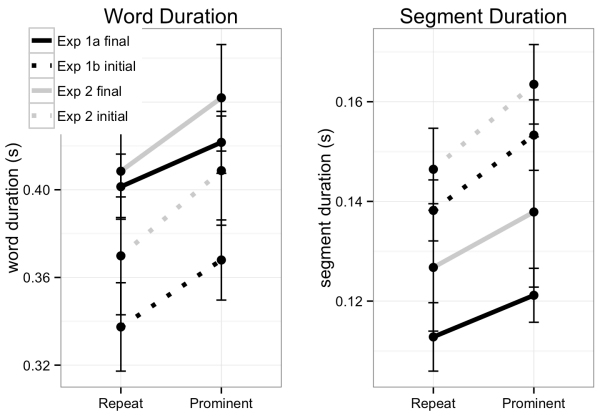
3.1. Prominence measures

Words in the *Prominent* condition were longer, louder and had higher f0 than words in the *Repeat* condition (Figure 2; $p<0.01$) indicating that they were produced with greater prominence [12, 21]. Sibilant segments in the *Prominent* condition were also longer and louder than in the *Repeat* condition ($p<0.01$). Both words and segments were shorter in Experiment 1a than the other two versions ($p<0.05$). There were no significant interactions between condition and experiment for words or segments.

3.2. Spectral distinctiveness - COG

Figure 3 shows the average COG values for each of the sibilants under the *Repeat* and *Prominent* conditions for Experiment 1 and 2. The interaction between segment and condition varied significantly by experiment (Exp1 vs. 2, $p=0.03$; Exp1a vs. b, $p=0.02$). An analysis on Experiment 1a and b found no change in segmental distinctiveness with prominence (condition*segment type, $p=0.17$) though there was a trend towards increased distinctiveness for Experiment 1b – onset segments (condition*segment*experiment, $p=0.06$). An analysis on Experiment 2 including segment position found increased segmental distinctiveness under prominence (condition*segment type, $p=0.02$) and no effect of onset versus coda segments.

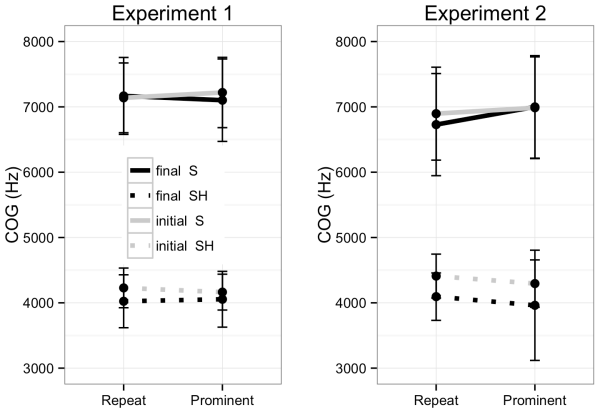
Figure 2: Durations of words and sibilant segments by condition, experiment and position of segment in the word. Error bars in all figures are standard error by participant.



3.3. Voice Quality – H1-H2

Figure 4 shows the average H1-H2 in preceding vowels by condition for the two experiments. Vowels were significantly breathier under prominence ($p<.01$) but this did not depend on experiment or segment.

Figure 3: Spectral Centre of Gravity (COG) by sibilant, condition, experiment and word position. Sibilants in Experiment 2 are more distinct in the *Prominent* condition where they contrast with a minimal pair.

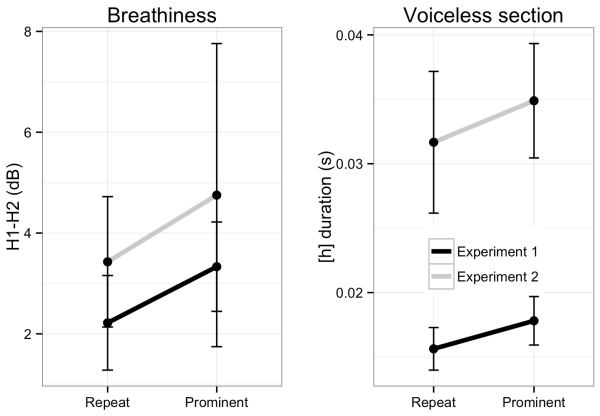


3.4. Voiceless vowel portions

An example of a voiceless vowel portion is given in Figure 1. Figure 3 shows the average duration of the voiceless vowel portion by condition and experiment. Voiceless portions were longer for Experiment 2 ($p<0.001$). A trend towards shorter sections before /s/ was not significant. Voiceless portions were slightly but not significantly longer for the *Prominent* condition ($p=0.09$) and there were more voiceless portions for the *Prominent* (128/265

tokens = 48%) than the *Repeat* condition (102/262 tokens = 39%).

Figure 4: Differences in voice quality and duration of the voiceless section by condition and experiment showing an increase in glottal opening under prominence and when there was a contrast with a minimal pair (Experiment 2).



3. DISCUSSION

We found that words in the *Prominent* condition – words that were in contrastive focus with the context sentence – were longer, louder and had higher f0 than words in the *Repeat* condition, indicating that they were more prominent. When we looked at the production of the sibilants however, despite segments being longer in the *Prominent* condition, we did not find that prominence alone made the sibilants more distinct from each other in terms of COG (no interaction between segment and condition in Experiment 1). This was not simply a limit on how carefully speakers could produce sibilants (i.e. not a ceiling effect) as we did find a modulation of distinctiveness in Experiment 2 when the prominent word contrasted with a minimal pair in the context sentence. These findings suggest that prominence alone is not enough to modulate segmental distinctiveness of the sibilant place contrast, but that under careful speaking conditions (minimal pair is present) speakers can produce the contrast more distinctively.

A second striking finding was that regardless of the presence or absence of a minimal pair, vowels preceding a sibilant were breathier and had longer and more voiceless portions in the prominent condition. This did not serve to increase the distinctiveness of the place contrast as unlike [25] we didn't find consistent differences in any measure according to which segment was produced. Also note that the increased breathiness in the preceding vowel is not a case of increased overlap of gestures between the vowel and sibilant due to casual speaking conditions, but increased gestural overlap due to more careful speaking conditions. This could perhaps be argued to serve the purpose of making the sibilant voicing contrast more distinctive. However, in the case of /ʃ/ at least the voicing contrast is limited to a few lexical items. Furthermore, the increased breathiness was even greater when talkers were contrasting with a minimal pair that differed in place. The alternative hypothesis we argue for here is that this is a case of prosodic strengthening as described by [18], here extended to sibilants in non-initial position.

While it is sometimes assumed that there is a direct relationship between prosodic strength and distinctiveness of acoustic contrasts, our results and those of the previous literature do not support this assumption. The data instead are consistent with prosodic modulation acting to strengthen (and lengthen) articulatory constrictions and glottal openings, which may sometimes also enhance acoustic contrasts. While we found evidence that talkers can modulate acoustic contrasts directly under specific conditions, previous results provide mixed results on when this occurs (e.g. see [26] for enhancement of durational contrasts only in careful speech, and only when correcting an apparent minimal pair error on the part of the listener; [3] for enhancement of cues to affect but not contrasts in infant directed speech; [5] for vowel space expansion that is not related to vowel inventory size and [16] for a reduction in vowel contrast with careful speech). Thus the weaker claim that acoustic contrasts are directly modulated by clear speech only [23], may also be incorrect. Rather, a general process of articulatory strengthening which may or may not enhance a particular acoustic contrast may be the norm and the direct enhancement of acoustic contrasts may in fact be rare.

Table 2: Target words from all three experiments.

Exp 1a
niece, niche, gas, gash, lass, lash, lease, leash, mass, mash, brass, cash, grass, crash, class, flash, moss, wash, glass, sash, miss, wish, geese, quiche, gloss, posh, boss, frosh, fuss, brush, bliss, fish
Exp 1b
socks, shocks, sought, shot, sop, shop, sack, shack, sag, shag, seep, sheep, seat, sheet, sip, ship, sin, shin
Exp 2
moss, mosh, sass, sash, gas, gash, class, clash, lass, lash, brass, brash, niece, niche, lease, leash, diss, dish, socks, shocks, sought, shot, sop, shop, sack, shack, sag, shag, seep, sheep, seat, sheet, sip, ship, sin, shin

7. REFERENCES

[1] Aylett, M., & Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America*, *119*(5), 3048-3058.

[2] Bates, D., Maechler, M., Bolker, B., Walker, S. 2014. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 3.1.2.

[3] Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, *36*(4), 847-862.

[4] Boersma, P., Weenink, D. 2013. Praat: doing phonetics by computer (Version 5.3. 57). University of Amsterdam. http://www.praat.org/.

[5] Bradlow, A. R. (2002). Confluent talker-and listener-oriented forces in clear speech production. *Laboratory phonology*, *7*, 241-273.

[6] Chang, Y. H. S., & Shih, C. (2012). Using Map Tasks to Investigate the Effect of Contrastive Focus on the Mandarin Alveolar-Retroflex Contrast. *Proceedings of Speech Prosody 2012*.

[7] Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of/ɑ, i/in English. *The Journal of the Acoustical Society of America*, *117*(6), 3867-3878.

[8] Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*(2), 155-190.

[9] Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics*, *39*(3), 344-361.

[10] Chuang, Y. Y., & Fon, J. (2010). The effect of prosodic prominence on the realizations of voiceless dental and retroflex sibilants in Taiwan Mandarin spontaneous speech. *Proceedings of Speech Prosody 2010*.

[11] Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics*, *35*(2), 180-209.

[12] Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question–answer contexts. *The Journal of the Acoustical Society of America*, *77*(6), 2142-2156.

[13] Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of phonetics*, *29*(2), 109-135.

[14] Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America*, *101*(6), 3728-3740.

[15] Gorman, K., Howell, J., & Wagner, M. (2011). Prosodylab-aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics*, *39*(3), 192-193.

[16] Granlund, S., Hazan, V., & Baker, R. (2012). An acoustic–phonetic comparison of the clear speaking styles of Finnish–English late bilinguals. Journal of Phonetics, 40(3), 509-520.

[17] Kang, K. H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, *124*(6), 3909-3917.

[18] Keating, P. A. (2006). Phonetic encoding of prosodic structure. Speech production: Models, phonetic processes, and techniques, 167-186.

[19] Keating, P., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, 157-173.

[20] Klatt, D. H., Stevens, K. N., & Mead, J. (1968). Studies of articulatory activity and airflow during speech. *Annals of the New York Academy of Sciences*, *155*(1), 42-55.

[21] Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, *118*(2), 1038-1054.

[22] Ladefoged, P. 1993. *A Course in Phonetics.* 3rd edn. San Diego: Harcourt Brace Jovanovich

[23] Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In: W. J. Hardcastle, & A. Marchal (Eds.), Speech production and speech modeling (pp. 403–439). Amsterdam: Kluwer Academic.

[24] Moon, S. J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical society of America*, *96*(1), 40-55.

[25] Niebuhr, O., & Meunier, C. (2011). The phonetic manifestation of French /s#ʃ/ and /ʃ#s/ sequences in different vowel contexts—on the occurrence and the domain of sibilant assimilation. *Phonetica*, *68*(3), 133-160.

[26] Schertz, J. (2013). Exaggeration of featural contrasts in clarifications of misheard speech in English.*Journal of Phonetics*, *41*(3), 249-263.

[27] Shadle, C. H., & Scully, C. (1995). An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences. *Journal of phonetics*, *23*(1), 53-66.