AUTOMATIC ANALYSIS OF SIBILANT ASSIMILATION IN ENGLISH

Clayards, Meghan[12], and Doty, Eric[2]
[1]School of Communication Sciences and Disorder, McGill University, 1266 Pine Ave. West, Montreal, Quebec, Canada
[2]Dept. of Linguistics, McGill University, 1085 Ave Dr. Penfield. Montreal, Quebec, Canada

1. INTRODUCTION

Assimilation across word boundaries is a common phenomenon in language. For example, the English coronal sibilant /s/ is pronounced more like a post alveolar sibilant /ʃ/ when followed by the post-alveolar sibilant (e.g. “glass shoe” pronounced as “glash shoe”).

Of theoretical interest in the extent to which these assimilation processes are complete versus partial and the extent to which they are under the cognitive control of the speaker. Holst and Nolan (1995) investigated sibilant assimilation in British English speakers and found a range of assimilation strengths from partial to complete. Niebuhr et al (in press 2011) found mostly cases of complete assimilation. In both studies the assimilation was directionally asymmetric in that /s/ assimilated to /ʃ/ when the /ʃ/ followed (e.g. “glass shoe”) but not when it preceded (e.g. “fish soup”). In contrast Neibuhr et al (in press 2011) found less directional asymmetry and less complete assimilations in the productions of French speakers. The differences between the French and English speakers and the (nearly) complete nature of the assimilations suggest that these processes are language specific and under the control of the speaker. Others have argued that even co-articulation (which may be equivalent to partial assimilation) may be under the control of the speaker. Whalen (1990) asked talkers to begin speaking before they knew the full VCV sequence He found co-articulatory effects of the consonant or second vowel on the first vowel *only* when their identity was known before hand. The lack of these co-articulatory effects in the absence of pre-planning strongly suggest that they are the result of pre-planning.

The purpose of the current study was to 1) test new automated methods for doing phonetic analysis by replicating previous work, 2) to compare sequences of identical sibilants as well as single sibilants (which was not done in Niebuhr et al) and 3) to investigate issues of planning of assimilation across word boundaries.

2. METHODS

Nineteen North American English speakers were recorded reading a set of sentences from a computer screen in a casual speaking. Recordings were made into a headset microphone and recording and stimulus presentation were controlled by Matlab.

2.1. Stimuli

All sentences were of the form, "Say X Y please/now," where words of type X and Y were nonsense words. In type X words one of the test phonemes /s/ or /ʃ/, or the control /p/ was at the right edge of the word and in type Y words at the left edge, resulting in sibilant sequences at the word boundary (e.g. /ʃ/-/s/ sequence in "Say caveesh sival please"). Each of the 3 consonants was used in two different nonsense words (e.g. /ʃ/ in "tamash" and "caveesh") resulting in 6 words for each of type X and Y. All 12 Words are listed in Table 1. These words were stimuli in a related perception experiment (Clayards, Gaskell, & Niebuhr, 2011). Participants were presented with all 36 X/Y combinations in a random order.

Table 1. Nonsense words used in production.

Word 1 (X)	tamash, caveesh, pidas, cavees, nalip, remope
Word 2 (Y)	shinnow, shamal, sival, samal, pentuf, pagoon

2.2. Analysis

Segmentation of the recordings was done automatically through forced alignment using HTK and the Prosody Lab aligner (Gorman, Howell, & Wagner, this volume). Figure 1 shows an example alignment with the accompanying spectrogram and waveform.

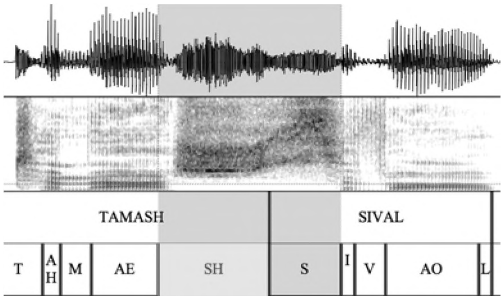

Figure 1. Example alignment displayed in Praat.

In many cases the segmentation of the sibilants included part of an adjacent vowel as in Figure 1. For this reason the first and last 20% of each sibilant (or sequence) was excluded from analysis. Segmentation of the sibilants within the sequence was not reliable, since in many cases there was no clear boundary between sibilants. For this reason, all analyses were done on the entire sibilant sequence.

A total of 603 productions were collected. Of these, 224 (37%) had a short silence between the two critical words. Productions without silence were analyzed separately. All analyses were conducted in Praat (Boersma & Weenik, 2010). Sibilant sequences were extracted and band-pass filtered between 1500 and 15000 Hz. Spectral slices were extracted every 7 ms using a 30ms Gaussian window and 50 Hz bins. Spectral centre of gravity (CoG) was calculated on each slice and the mean and range of the CoG measurements was calculated for each sibilant sequence.

3. RESULTS

Figure 2 shows the distribution of mean CoGs for each of the single sibilants and mixed sibilant sequences investigated by Niebuhr et al. (in press, 2011).

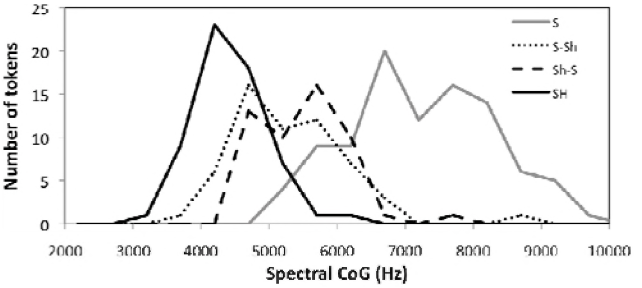

Figure 2. Distribution of mean CoG for single sibilants and sibilant sequences.

The CoG of both of the mixed sibilants are on average closer to /ʃ/ than to /s/ indicating some assimilation and the /s ʃ/ sibilants are most like /ʃ/. This pattern is much more like Holst and Nolan (1992) than Niebuhr et al. (in press, 2011) in that the /s ʃ/ sequences show a range of assimilation strengths and most are not complete.

The second goal was to make comparisons between the single sibilants (in the context of /p/) and /s s/ and /ʃ ʃ/ sequences. Table 2 lists the means and ranges (as well as standard deviations) of each of the sibilant conditions.

Table 2. Mean and range of CoG (SD) for each sibilant.

Sibilant(s)	Mean CoG (Hz)	CoG range (Hz)
/s s/	7333 (1041)	3224 (1827)
/s ʃ/	5040 (849)	3000 (1663)
/ʃ s/	5245 (609)	4203 (1569)
/ʃ ʃ/	4083 (523)	1524 (731)
/s/	6999 (1049)	2504 (1435)
/ʃ/	4189 (537)	1048 (747)

The mean CoG of the /s s/ sequence is 334 Hz higher than the singleton /s/ and this difference is marginally significant (t = 1.94, df = 126.1, p = 0.054). The /ʃ ʃ/ sequence is 106 Hz lower than the /ʃ/ singleton but this difference is not significant (t = -1.07, df = 112.6, p = 0.286).

Finally, we examined the relationship between planning and assimilation by analysing the 224 productions in which a short pause had been inserted between the sibilants. Here we found very little assimilation. The /s/ sibilants produced before /ʃ/ were nearly identical to those in other contexts. We did however, find that there was a significant correlation between the length of the silence (log ms) and the CoG of /s/ produced before /ʃ/ (R^2 = 0.21, t = 2.13, df = 17, p(two tailed) = .048). This is illustrated in Figure 3. No such relationships existed between any of the other segments and pause duration.

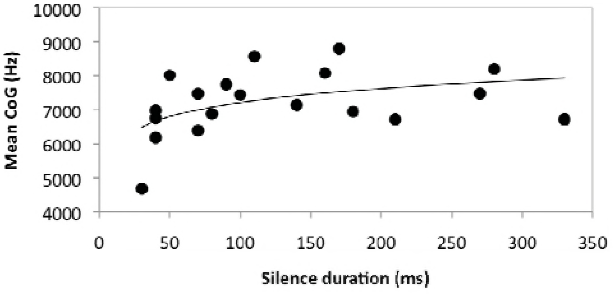

Figure 3. Relationship between duration of the silence and mean CoG of /s/ before /ʃ/.

4. DISCUSSION AND CONCLUSIONS

Our results are in line with previous studies, using much less labour intensive methods. However, discrepancies will need to be investigated further to determine if they are due to analysis techniques or dialect differences. Single /s/ in the context of a labial consonant had a significantly lower mean CoG than /s s/ sequences. Finally, assimilation is restricted to cases where there is no pause between words, but traces of co-articulation are visible for short pauses suggesting they could be due to the degree of articulatory planning.

REFERENCES

Boersma, P., & Weenik, D. (2010). Praat: doing phonetics by computer (Version 5.1.31). 2010

Clayards, M., Gaskell, G., & Niebuhr, O. (2011). *Comparing French and English listeners' on-line perception of assimilated speech.* Paper presented at the BCBL 10th International Symposium on Psycholinguistics, Donostia-San Sebastian.

Gorman, K., Howell, J., & Wagner, M. (this volume). Prosodylab-Aligner: A tool for forced alignment of laboratory speech. . *Proceedings of Acoustics Week, the annual conference of the Canadian Acoustical Association.*

Holst, T., & Nolan, F. (1995). The influence of syntactic structure on [s] to [ʃ] assimilation. In B. Connell & A. Arvaniti (Eds.), *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence* (pp. 315-333). Cambridge: CUP.

Niebuhr, O., Clayards, M., Lancia, L., & Meunier, C. (in press 2011). On place assimilation in sibilant sequences - comparing French and English. *Journal of Phonetics.*

Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics, 18*, 3-35.