LIN 825

GESTURAL OVERLAP, PERCEPTABILITY, AND SONORITY

February 7, 2020

Isaac Sarver
Michigan State University
Department of Linguistics and Languages
East Lansing, Michigan, USA
sarveris@msu.edu

ABSTRACT

This paper evaluates prior work on the relationship between gestural overlap as measured by articulatory data and recoverability. Data and claims made by Chitoran et al. (2001) provide the groundwork based on evidence from Georgian that the recoverability of segments plays a role in determining their sensitivity to overlap in certain contexts. If this is true, then a definition of sonority sequencing based on recoverability links the sonority to the overlap facts. Predictions of this work are discussed, and then investigated through a study of articulatory data taken from the Wisconsin X-Ray Microbeam Database corpus. This study suggests that while the Georgian data is not comprehensive evidence, and the English data from the corpus is also limiting in what it can show, the results of this paper are consistent with the current story, and can point to future areas of inquiry.

1 INTRODUCTION

Articulatory data has been used in many ways to try and illuminate the interface between phonetics and phonology. The timing and shape of gestures can give information about the timing of phonological cues and triggers that would otherwise be lost in the acoustic counterpart of the data.

Just as in the acoustics, the articulatory patterns of speech are not clear, linear, discrete segments. Whatever the phonology is assumed to be, there are processes which take the phonological information and translate that into the phonetics. How that happens, and how speech could translate from a traditional surface representation in the phonology to the gestures and sounds that make up the phonetics and speech is not entirely known. However, looking at the patterns that articulatory data can provide may give insights into this process and illuminate the parameters that govern how these gestures can occur.

In this paper, I will begin with a discussion of prior work on articulatory data, looking at Chitoran et al. (2001) in Section 2. This study of overlap in stop-stop sequences in Georgian raises questions for further work. I will look at their predictions and results and motivate the questions and predictions made in my study. In Section 3, I will discuss the articulatory data I will be using for English and what aspects of the data will be used for the analysis. In Section 4, I will discuss the results that can be shown from that data, and in Section 5, I will offer discussion, implications, and conclusions of the results.

2 PRIOR WORK

2.1 Predictions

One premise of the prior work on gestural overlap concerning CC clusters is that this gestural overlap can be restricted or allowed based on the ease of perceptability. If there is a need for high perceptability in the potential contexts of the segments, they will exhibit low overlap, and if this requirement of high perceptability is removed, then a higher degree of overlap between gestures will be allowed. Chitoran et al. make the case that this is apparent when two stops are in a word-initial onset (following Redford and Diehl, 1999). Word-initial onsets are potential utterance onsets, and if a stop is at the onset of an utterance, there is a lack of acoustic information preceding the release of the stop. Because there is no preceding vowel, there are no formant transitions from a vowel into the stop. In languages that allow stop-stop onsets, such as Georgian, this is true for both the C1 and the C2 of the onset.

Because of this, two successive stops which fall across a word boundary should not have as much restriction on the amount of overlap. Acoustic studies have been done to study this effect in Georgian, but there is not a consistently clear way to test for overlap in the acoustics. The method used in that study (Chitoran 1999) was to test for release burst of C1, which, if not present, was tied to a large amount of overlap obscuring the release. This is also not a good way to judge the effect, as there could be other factors affecting stop release word-medially and word-finally.

Based on these premises, the paper tests whether it is true that word-initial stop-stop sequences have less overlap than word-internal stop sequences. They also tested for effects of the order of the stops; whether back-to-front gestures (coronal-labial, dorsal-labial, dorsal-coronal) exhibit more or less overlap than the front-to-back gestures. As will later be discussed in Section 3, these effects are much harder to test for with the articulatory data available to me, and unless relevant to the discussion, I will focus primarily on the implications of the position of the sequences rather than the order.

2.2 Experiments on Georgian Stop-Stop Sequences

The study on Georgian (Chitoran et al. 2001) was done using the EMMA (Electromagnetic Midsagittal Articulometer) magnometer system. Two Georgian speakers were asked to repeat the words in table 1. Details of the methodology used for collecting data via EMMA will be discussed in Section 3.

For the analysis of the results, the study used a 3-factor full-interaction ANOVA with overlap

consonants C1	C2	Word-Initial Sequences		Word-Internal Sequences	
front-to-back					
b	g	bgera	'sound'	abga	'saddle bag'
p^h	t^h	p^ht^hila	'hair lock'	ap^ht^har-i	'hyena'
d	g	dg-eb-a	's/he stands up'	a-dg-eb-a	's/he will stand up'
back-to-front					
g	b	g-ber-av-s	's/he is inflating you'	da-gbera	'to say the sounds'
t^h	b	t^hb-eb-a	'it is warming up'	ga-t^hb-a	'it has become warm'
g	d	gd-eb-a	'to be thrown'	a-gd-eb-a	'to throw sth. in the air'

Table 1: Stimuli for Georgian experiment

as the dependent variable for each speaker separately. The results were as follows:

Effect	Speaker 1	Speaker 2
POSITION	**F(1,69)=51.48, p<.001**	F(1,68)=.06, p>.05
ORDER	**F(1,69)=52.91, p<.001**	**F(1,68)=8.37, p<.01**
PLACES	**F(2,69)=17.03, p<.001**	**F(2,68)=8.63, p<.001**
POSITION X ORDER	F(1,69)=.23, p>.05	**F(1,68)=8.07, p<.01**
POSITION X PLACES	F(2,69)=.81, p>.05	F(2,68)=3.41, p>.01
ORDER X PLACES	F(2,69)=.16, p>.05	**F(2,68)=7.95, p<.01**
POSITION X ORDER X PLACES	S F(2,69)=1.5, p>.05	F(2,68)=3.56, p>.01

Table 2: Summary of Chitoran et al. results (significant effects in bold)

There are three aspects of this study that I want to highlight as reasons for caution when interpreting these results. The first is that these results only find ORDER and PLACES as significant across both speakers. Both of these calculations rely on dorsal-coronal comparisons. It is not clear how trustworthy the gestural data of these two gestures can be. Both are tongue gestures, and are thus not independent. It should not be assumed that there are no mechanical accommodations made when there are adjacent dorsal and coronal gestures. When considering something like whether the order is front-to-back or back-to-front, it is not safe to assume that bg/gb comparison will operate under the same constraints as a dg/gd comparison.

The second is that comparisons are made between voiceless aspirated consonants and voiced unaspirated consonants. There is a lack of an account for potential allophony or rules which could change amounts of aspiration of segments in different positions, and this could have an effect on the results. Additionally, the coronal-labial comparison is made between p^h/t^h and t^h/b, so the study did not use minimal pairs in these cases.

The third is that the study conducts ANOVAs on individual speakers. This is problematic because the independence assumption for the ANOVAs is thrown into question. The fact that the study only used two speakers is limiting, and there could be much more gained from an across-speakers comparison. Even looking at table 2, the difference in significant results between the two speakers adds to the caution that should be taken.

2.3 Perceptability and Sonority Sequencing

The results do however have implications for sonority sequencing violations. All of the data in the Georgian example were examples of sonority plateaus, or stop-stop sequences, which are sonority-violating and cross-linguistically uncommon. Sonority sequencing usually just refers to the cross-linguistic patterns of consonant ordering towards and away from the nucleus of a syllable, with the onset having rising or increasing sonority in most languages and the coda having falling sonority. This leads to a sonority scale from obstruents to vowels:

(1) Obstruents < Nasals < Liquids < Glides < Vowels

Manner is also often further specified in this scale so that stops precede fricatives. The premise of the Chitoran et. al study was based upon motivations from perceptability or recoverability of the segments by the listener. Crucially, arguments are also made for perceptability being a phonetic correlate of the sonority scale (Mattingly, 1981; Ohala, 1990). The claim is that there is a correlation between the least sonorous segments and the least easily perceptible segments, as well as between the most sonorous segments and the most easily perceptible segments. Just as discussed earlier concerning utterance-initial stops, obstruents need the most protection against loss in perceptability, as opposed to nasals, liquids, vowels etc.

If this is the case, then one place this should be apparent is the common sonority-violating sC onsets. These are present in English as /sk/, /st/, /sp/, and others, and since there is a fricative followed by a stop, they are described as falling-sonority, and should have an effect similar or greater to that of the sonority-plateau stop-stop sequences in Georgian. The subsequent sections I will use to discuss my own findings when investigating these implications in English using articulatory data as well.

3 DESIGN

3.1 The data

In order to investigate the prediction that sonority sequencing can have an effect upon gestural overlap, I will use data from the Wisconsin X-ray Microbeam Corpus. This corpus is a collection of speech production data from native English speakers reading a collection of word lists, sentences, paragraphs, and performing other speech tasks.

The gestural data is collected via a number of sensors which are placed on the subjects' tongue, lips and head. These sensors can be detected through x-rays. Anchor sensors on the subject's head can be used to measure the relative movement of the parts of the tongue and the lips as the subject speaks. This data can be displayed alongside acoustic data, with the x-axis being time in milliseconds and the y-axis being the measurements collected during speech.

The gestures in question, in for instance an /sp/ onset in the word 'special,' being produced by a participant reading from a word list, would be comprised of two gestures in the data. The tongue-tip gesture for the /s/ segment, and the labial gesture for the /p/ segment. These gestures can be identified in the data visually using MATLAB. After the gestures are found, the 'findgest' algorithm is used in MATLAB to find the necessary measurements. The algorithm is activated by clicking on a point in a sensor's data displayed linearly, say the tongue-tip sensor. The algorithm then takes the surrounding values, and from the local min and max values calculates a movement and assigns a series of points along that movement such as gesture onset, plateau onset, plateau offset and gesture offset.

In general, data from 26 of the speakers was used for all of the results discussed in section 4. However, due to some inconsistencies in the recording, some speakers were eliminated. If an identical environment to the one in question was available in another file, than this data was used instead. (For example, the word 'special' is used many times in word lists across different files throughout the corpus.) For some sequences though, there were not other instances in the corpus, and if there was an issue with a speaker's recording of the one instance needed, their data was lost from the study.

For this paper, I only collected the gesture onset and offset data. This can be used to calculate the overlap of the two gestures, which I define in the following way:

(2) OVERLAP - The overlap of segments x and y is $x_{offset} - y_{onset}$.

There are a few issues with the corpus data that relate to the questions at hand. The data was not collected with these questions in mind. The word lists, sentences, and paragraphs produced by the speakers for the corpus are not catered to the comparisons I wish to make. Ideally, a new experiment could test for all the possible environments of interest for this paper, but for now I believe the crucial insights that can be gained from the Wisconsin Microbeam Database have been examined.

3.2 Comparisons

This study checks three comparisons in the data. The first is the comparison between a position. Two segments (/s/ and /m/) combined in the onset of a word and across a word boundary. This is in part to try and see if the prediction of the Chitoran et al. claims bear out. This comparison cannot claim to show much; English does not have stop-stop sequences as in Georgian, and it is harder to find onsets where these effects may be observed. Even though there might be effects in a comparison of this sort between /s/ and /p/, the confound with that comparison is that there is simply no data in the corpus to use for that comparison. The same goes for /s/ and /w/, and any other comparison that is both useful and trustworthy. As discussed in Section 2, using dorsal-coronal comparisons can be problematic. As such, the comparison which is left is the /s#m/ vs /sm/ case. This is then predicted to not show any overlap sensitivity in the onset because since the onset is sonorous, it should not have recoverability issues that prompt overlap sensitivity.

The second comparison is to compare length of a gesture with the difference in gesture onsets. The goal of this comparison is to use the gesture length as a proxy for speech rate. This comparison should show whether the two segments are really just one complex segment. If the first comparison shows evidence for overlap sensitivity, the story will be clearer if these segments can be shown to be distinct, otherwise the overlap data may just be a result of the fixed complex segment in the phonology and the overlap patterns happen to be a coincidence, and not produced by the interaction of two separate phonological elements as they are transferred into the phonetics.

The third is the comparison between a rising-sonority onset and falling-sonority onset. This data is available in the corpus in a word list, where the words ‘smooth’ and ‘special’ are produced. According to the predictions stated regarding sonority and overlap in Section 2, the prediction for this data is that /sm/ will have a higher degree of overlap than /sp/, since /sp/ is a falling-sonority onset and thus is more sensitive to overlap.

4 RESULTS

The results of this study showed a statistically significant effect of overlap on sonority sequencing, a correlation between speech rate and onset difference times, and no significant effect of position (for /s#m/ and /sm/) on overlap. The section will present the data visualized in a few ways and explain each case. For the data visualizations, the tidyverse and ggplot packages in R were used.

4.1 Position

The boxplot in figure 1 shows the comparison between amount of overlap in milliseconds for tongue-tip and labial gestures depending on whether the context is across word boundaries or in an onset:

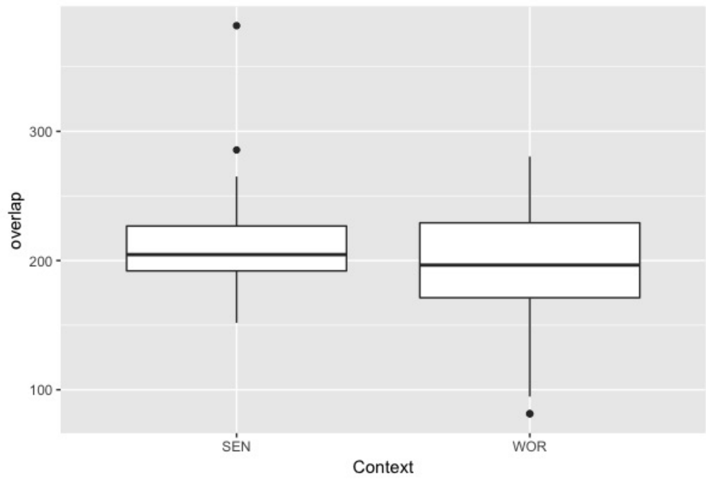

Figure 1: SEN=/s#m/ and WOR=/sm/

These examples were taken from a paragraph where the speaker said 'this March' and 'small' in a sentence near the end of the recording. There is no significant effect of position on the amount of overlap between these segments [t = 1.26, p > 0.05]. Though both contexts have edge values that skew in one direction vs. the other, these are not enough to give any significant result. What is actually impressive is how similar the overlap is in these examples. This is especially clear when viewed in a density plot. See figure 2.

One way to examine this comparison further is to not look at the overlap in seconds, but rather the percentage overlap per gesture. The reason for this is that maybe there is a regulated way that the length of the gestures is changing in the different contexts, and though the length of the overlap might not change, its ratio to the gesture as a whole might.

This can be tested with the current data, with percentage overlap calculated as the previous overlap calculation divided by the length of the labial gesture (Labial offset − labial onset). Doing this does not give a significant result either ([t=−1.63, p>0.05]).

As discussed beforehand, this comparison cannot show the predictions for falling-sonority sequences in these contexts. Though, since it is the only available comparison to be found in the corpus, it is still useful to know that it falls in line with that story, and that since the /sm/ case is a sonorous sequence, there is expectation for a difference in overlap depending on the contexts discussed here.

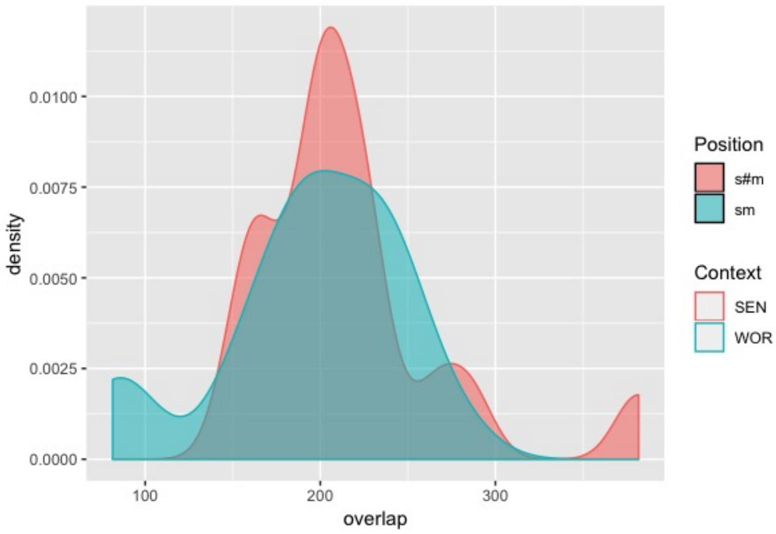

Figure 2: Comparing Overlap in Milliseconds

4.2 Speech Rate

Comparing the length of the sequence to the difference in onsets of the two gestures shows that they are correlated. If the length of the sequence increases, the difference between the onsets also grows. See figure 3 and 4.

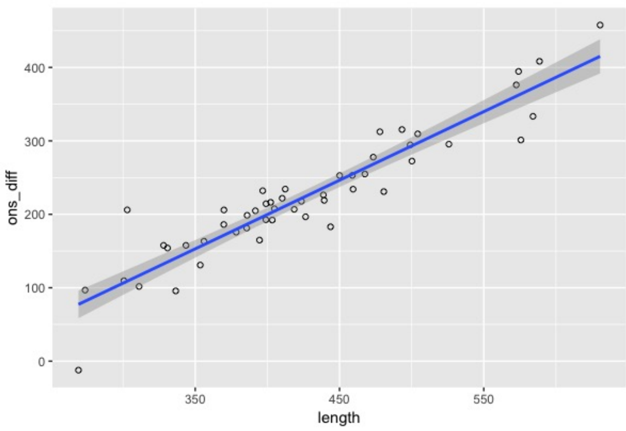

Figure 3: Correlation between Sequence Length and Difference of Onsets

These comparisons support an important assumption of the study, which is that onsets like /sm/ and /sp/ are complex onsets made up of multiple segments, and not complex segments. If /sp/ were instead a complex segment, there would be little to be said about the phonetic realization of overlap between the gestures of this segment being a result of a compositional

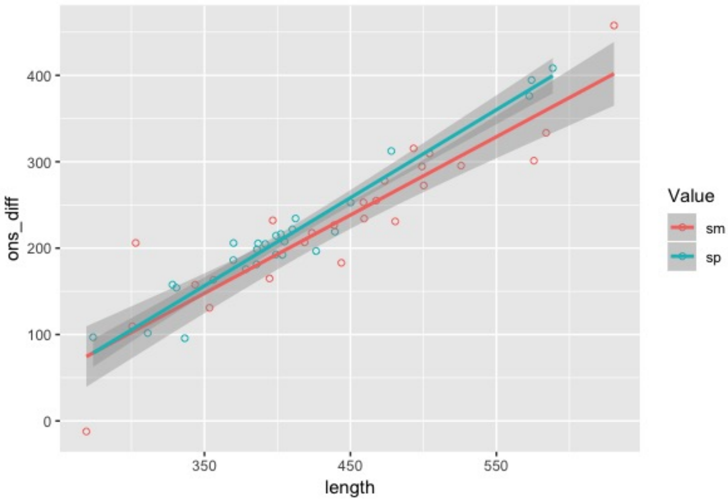

Figure 4: Speech Rate Comparison for both /sm/ and /sp/

process, taking into account the sonority and need for recoverability. Rather, the overlap should not be assumed to be a fact of 'sensitivity' when instead it could be a setting specified in the segment representation.

4.3 Sonority Differences: /sm/ and /sp/

The third comparison proved to be the most visually rewarding result given the predictions of Section 2. The comparison here is simply a test of whether the overlap of /sp/ is significantly more restricted than the overlap of /sm/ in onsets. Since /sp/ is falling-sonority, this should be the case. As figure 5 and 6 show, /sm/ tends to have much more overlap.

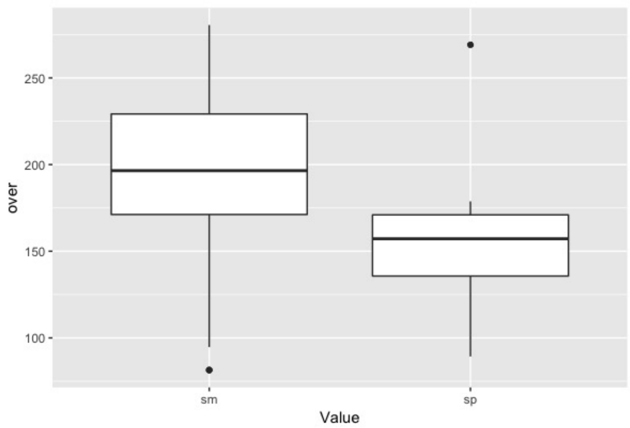

Figure 5: Difference in Overlap for /sm/ and /sp/ Onsets

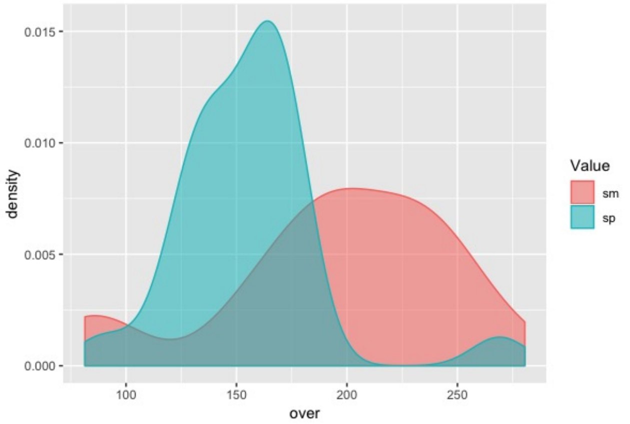

Figure 6: Density Plot of Overlap Differences

These results show a significant effect, as running a Welch test shows: [T = 3.227, P = 0.0025].

5 Conclusion

Given the results in the previous section, I propose that the English data in the Wisconsin X-Ray Microbeam Database corpus follows the patterns predicted based on prior work by Chitoran et al. and are consistent to the extent that can be shown. The results for the position of /s/ and /m/ relative to their overlap can be used to show that since stops in Georgian can have significantly different overlap in these positions, the /s/ and /m/ comparison can act as a sort of control case for that claim. The speech rate results also remain consistent with this, and the sonority difference comparisons do bear out the way that is predicted.

However, much of the caution I spoke of in section 2 regarding the results of the Georgian data is in no way mitigated by these results. They do not suffer from the same complications, but there are some concerns that should be raised.

Though these comparisons are made across a larger group of speakers, I was very limited in the choice I had for segments to compare. This is both due to the corpus itself, and the data within, as well as due to the phonology of English. The sC and sCC onsets are the only falling-sonority sequences in English. sC is often singled out cross-linguistically as a special case of sonority violation, being allowed very often when no other sonority-violating sequences are. The fact that no other types of falling-sonority sequences can be tested for in English is one issue.

Secondly, comparing /sm/ to /sp/ does not leave out the possibility that maybe the increase in overlap with the /sm/ onsets is due to another fact, such as the nasal. If it were possible to test the gestural data for another sonorous sC sequence, like /sl/, that could be used as a control, but the fault of the articulatory cues is that very few sequences can be separated into clear, independent gestures in the articulatory data.

Articulatory data is valuable for analysis of the phonetics-phonology interface, and that should not be understated. The current study is limited in the data available, but any future study will also be limited by the nature of what articulatory data can show. Nevertheless, future experiments could bear quite strongly on this story of gestural overlap and perceptibility. A chance to gather more data with EMMA could lead to a more thorough recreation of the Chitoran et al. experiment with more speakers and more careful statistics, and could be conducted on Georgian, Russian, and other languages where stop-stop sequences and other crucial sequences could be analyzed.

REFERENCES

[1] Chitoran, I. 1999. ‘Accounting for sonority violations: the case of Georgian consonant sequencing.’ In *Proceedings of ICPhS XIV, San Francisco, 1999, 101-104.*

[2] Chitoran, I., Goldstein, L. & Bird, D. 2001. ‘Gestural Overlap and Recoverability: Articulatory Evidence from Georgian’. In *C. Gussenhoven, T. Rietveld, and N. Warner (eds.) Papers in Laboratory Phonology 7, Cambridge University Press.*

[3] Mattingly, I.G. 1981. ‘ Phonetic representation and speech synthesis by rule.’ In *The Cognitive Representation of Speech (T. Myers, J. Laver and J. Anderson, editors), pp. 415-420. North Holland Publishing Company.*

[4] Ohala, J.J. 1990. ‘Alternatives to the sonority hierarchy for explaining segmental sequential constraints.’ In *Papers from the Chicago Linguistic Society 26: vol. 2, The parasession on the syllable in phonetics and phonology, 319-338.*

[5] Redford, M.A. and Diehl, R.L. 1999. ‘ The relative perceptual distinctiveness of initial and final consonants in CVC syllables.’ In *Journal of the Acoustical Society of America, 106, 1555-1565.*