2025. Proc Ling Soc Amer 10(1). 5916. https://doi.org/10.3765/plsa.v10i1.5916.

The CV coordination in English and Mandarin

Yunting Gu, Ben Airola, Yichen Wang, Anqi Shao & Karthik Durvasula*

Abstract. There is no consensus on the phonetic basis of sonority, either in the articulation or the perception of speech (Albert 2023). This study explores sonority in speech production by following up on previous studies which suggest a positive correlation between CV lag and sonority difference (Crouch 2022; Crouch et al. 2023; Gao 2008; Shaw & Chen 2019). Based on Electromagnetic Articulography (EMA) experiments participated by 10 English and 10 Mandarin speakers, we found that CV lag positively correlates with CV sonority difference in both languages. If we make an assumption that larger lags are preferred within a syllable, the finding forms a basis to explain universal constraints such as the Sonority Sequencing Principle (SSP) and the Sonority Dispersion Principle (SDP).

Keywords. sonority; EMA; speech production; SSP; SDP; CV lag

1. Introduction. Sonority has been one of the most debated concepts in phonetics and phonology. Constraints involving sonority such as the Sonority Sequencing Principle (SSP), the Sonority Dispersion Principle (SDP), or the Syllable Contact Law have long been used by phonologists to understand syllable structure (Sievers 1881, 1901; Steriade 1982; Selkirk 1984; Clements 1990; Kenstowicz 1994; Parker 2002, 2011). However, there is no consensus on the phonetic basis of sonority, either in the articulation or the perception of speech (Albert 2023). This study explores sonority in speech production.

Current speech production theories do not predict the variation of gestural coordination relevant to sonority. However, sonority has been observed to be correlated with systematic variation in gestural coordination, based on CC clusters in Georgian (Crouch 2022). Also, some observations (Gao 2008; Shaw & Chen 2019) suggest that sonority seems to be a factor that systematically correlates with CV gestural coordination variation. In this study, we followed up on these previous studies and explored whether there is a positive correlation between sonority difference and CV lag (the gestural lag between a consonant and a vowel) in English and Mandarin. There are two sub-claims to be tested as in (1).

(1) a. For CV syllables with the same V, a less sonorous C leads to a larger CV lag.
b. For CV syllables with the same C, a more sonorous V leads to a larger CV lag.

2. Methods.

2.1. MAIN VARIABLES IN THE STUDY. The study evaluated the relationship between *sonority difference* and *CV lag*. In order to quantify the dependent variable, namely, CV lag, we measured the timing difference between the consonant gesture and the following vowel gesture (CV lag = V timestamp - C timestamp). This way of calculating lag by subtracting corresponding timestamps was used in Zhang et al. (2019). Specifically, the CV lag in the current study was computed by subtracting the target onset (onset of gestural plateau) timestamp of the consonant from the target onset timestamp of the vowel. Target onset instead of gestural onset was used

* Authors: Yunting Gu, Michigan State University (yuntingguu@gmail.com) & Other Authors, Michigan State University.

© 2025 Author(s). Published by the LSA with permission of the author(s) under a CC BY license.

since target onset alignment has been argued to be more consistent than gestural onset alignment (Zhang et al. 2019; Durvasula & Wang 2023).

To quantify sonority difference, we considered many proposed sonority scales that are subtly different in the literature (Clements 1990; Kenstowicz 1994; Mielke 2008; Parker 2008; Kang et al. 2011). Ultimately, we chose to implement the sonority scale developed by Parker (2002, 2008, 2011) that is partially shown in Table 1. This scale was used since it is phonetically grounded and intended to cover all speech sound categories. While providing a much more nuanced sonority scale, its relative ranking of major classes accords with other sonority scales (Clements 1990; Kenstowicz 1994; Smolensky 1995; Clements 2005).

Natural class	**Sonority index**
low vowels	17
mid peripheral vowels (not [ə])	16
high peripheral vowels (not [ɨ])	15
mid interior vowels([ə])	14
laterals	9
nasals	7
voiced fricatives	6
voiced stops	4
voiceless fricatives (including [h])	3
voiceless stops (including [ʔ])	1

Table 1. Partial hierarchy of relative sonority (Parker 2002, 2008, 2011).

2.2. STIMULI. EMA data of 24 English stimuli and 26 Mandarin tone 4 stimuli were collected and analyzed to test the main claim. A summary of all English stimuli can be found in Table 2. The sonority difference of each stimulus can be found in the last column.

A summary of all 27 Mandarin stimuli can be found in Table 3. As can be seen in both Table 2 and Table 3, there are stimuli with coronal and (bi)labial consonants combined with high, mid, and low vowels.

2.3. PROCEDURES. Each set of stimuli in the EMA experiments was read 15 times in different randomized lists. This means that there were 15 repetitions of each stimulus. When collecting EMA data, sensors were glued to the tongue tip, tongue blade, tongue dorsum, upper lip, and lower lip of each participant. All the kinematic data were annotated in Matlab using the *lp_findgest* algorithm of the *mview* package (Tiede 2005), where the landmarks were labeled at 20 percent thresholds of peak velocity. Plots and mixed effects modeling were generated in R (R Core Team 2017) where CV lag was modeled as a function of the sonority difference, with participant, stimuli, and C duration as random intercepts.

3. Results. The study found a significant positive correlation between sonority and CV gestural coordination in English and Mandarin. For all the data, CV lag positively correlates with sonority difference significantly. Sub-groups of the stimuli controlled for consonant place of articulation or vowel height mostly exhibited the expected correlations. We also used consonant displacement and vowel displacement as estimates for jaw movement, and these findings suggest that jaw movement may not be a valid alternative account for the finding.

Index	Stimuli	C	V	C category	V category	Sonority diff
1	peak	p	i	bilabial	high	14
2	beak	b	i	bilabial	high	11
3	meek	m	i	bilabial	high	8
4	week	w	i	bilabial	high	3
5	pain	p	e	bilabial	mid	15
6	bane	b	e	bilabial	mid	12
7	main	m	e	bilabial	mid	9
8	wane	w	e	bilabial	mid	4
9	back	b	æ	bilabial	low	13
10	pack	p	æ	bilabial	low	16
11	Mac	m	æ	bilabial	low	10
12	whack	w	æ	bilabial	low	5
13	two	t	u	coronal	high	14
14	sue	s	u	coronal	high	12
15	do	d	u	coronal	high	11
16	new	n	u	coronal	high	8
17	toe	t	o	coronal	mid	15
18	so	s	o	coronal	mid	13
19	doe	d	o	coronal	mid	12
20	know	n	o	coronal	mid	9
21	talk	t	ɑ	coronal	low	16
22	sock	s	ɑ	coronal	low	14
23	dock	d	ɑ	coronal	low	13
24	knock	n	ɑ	coronal	low	10

Table 2. English stimuli summary.

Index	Word	Pinyin	T	C	V	C cat	V cat	Gloss	S dif
1	僻	pi	4	p	i	bilabial	high	distant	14
2	臂	bi	4	b	i	bilabial	high	arm	11
3	秘	mi	4	m	i	bilabial	high	secret	8
4	帕	pa	4	p	a	bilabial	low	napkin	16
5	坝	ba	4	b	a	bilabial	low	dam	13
6	骂	ma	4	m	a	bilabial	low	scold	10
7	袜	wa	4	w	a	bilabial	low	sock	5
8	配	pei	4	p	ə	bilabial	mid	match	13
9	贝	bei	4	b	ə	bilabial	mid	shell	10
10	妹	mei	4	m	ə	bilabial	mid	sister	7
11	味	wei	4	w	ə	bilabial	mid	flavor	2
12	肺	fei	4	f	ə	labial	mid	lung	11
13	盼	pan	4	p	æ	bilabial	low	hope	16
14	半	ban	4	b	æ	bilabial	low	half	13
15	曼	man	4	m	æ	bilabial	low	grace	10
16	万	wan	4	w	æ	bilabial	low	ten thousand	5
17	饭	fan	4	f	æ	labial	low	meal	14
18	兔	tu	4	t	u	coronal	high	rabbit	14
19	素	su	4	s	u	coronal	high	plain	12
20	度	du	4	d	u	coronal	high	degree	11
21	怒	nu	4	n	u	coronal	high	anger	8
22	路	lu	4	l	u	coronal	high	road	6
23	踏	ta	4	t	a	coronal	low	step	16
24	飒	sa	4	s	a	coronal	low	cool	14
25	大	da	4	d	a	coronal	low	big	13
26	那	na	4	n	a	coronal	low	that	10
27	腊	la	4	l	a	coronal	low	wax	8

Table 3. Summary of Mandarin stimuli. The consonant and vowel categories can be found in *C cat* and *V cat* columns. The sonority difference of each stimulus can be found in the last column *S diff*.

	Dataset (English EMA data)	Estimate (sonority difference)
	All English EMA data	11.24 ***
Bilabial C	All bilabial C data, C and V displacement	12.74 ***
	High V (week, meek, beak, peak)	14.00 ***
	Mid V (wane, main, bane, pain)	14.35 ***
	Low V (whack, Mac, back, pack)	10.47 ***
Coronal C	All coronal C data, C and V displacement	8.31 *
	High V (new, do, sue, two)	8.02 **
	Mid V (know, doe, so, toe)	9.40 ***
	Low V (knock, dock, sock, talk)	7.43 ***

Table 4. Summary of English EMA results. *** means that $p \leqslant 0.001$; ** means that $p \leqslant 0.01$; * means that $p \leqslant 0.05$.

The summary of the English experiment results can be found in Table 4 and Table 5. In Table 4, all the subgroups showed that there is a significant positive correlation between CV lag and sonority difference. The bilabial C group had a slightly larger effect size than the coronal C group.

	Pairwise comparison	Stimulus pair	Estimate (sonority difference)
Bilabial C	Nasality differ	Mac, back	11.20 *
		meek, beak	21.84 ***
		main, bane	21.59 ***
	Voicing differ	beak, peak	8.92
		bane, pain	12.81 **
		back, pack	12.66 *
	Vowel height differ	peak, pack	3.2
		beak, back	-0.22
		meek, Mac	7.61
		week, whack	18.23 **
Coronal C	Nasality differ	new, do	23.74 ***
		know, doe	11.83 **
		knock, dock	17.58 ***
	Voicing differ	two, do	-5.64
		toe, doe	6.47
		talk, dock	-0.23
	Vowel height differ	two, toe	61.90 ***
		do, doe	27.51
		sue, so	76.03 ***
		new, know	60.78 ***

Table 5. Summary of English EMA pairwise comparison. *** means that $p \leqslant 0.001$; ** means that $p \leqslant 0.01$; * means that $p \leqslant 0.05$.

In Table 5, the pairs which differ in nasality for both bilabial C and coronal C stimuli

	Dataset (Mandarin EMA data)	Est (son diff)	Est (displ)
	All Mandarin EMA data	7.11 **	
Bilabial C	C, V displacement as random intercepts	3.65	
	V displacement as fixed effect	5.73 **	11.90 ***
	High V (mi4, bi4, pi4)	10.23 ***	
	Mid V (wei4, mei4, bei4, pei4)	5.55 ***	
	Low V no coda (wa4, ma4, ba4, pa4)	5.73 ***	
	Low V with coda (wan4, man4, ban4, pan4)	5.61 ***	
Coronal C	C, V displacement as random intercepts	15.02	
	High V (lu4, nu4, du4, su4, tu4)	16.54 ***	
	Low V (la4, na4, da4, sa4, ta4)	19.84 ***	

Table 6. Summary of Mandarin EMA results. *Son diff* means sonority difference, *est* means estimate, and *displ* means displacement. *** means that $p \leqslant 0.001$; ** means that $p \leqslant 0.01$; * means that $p \leqslant 0.05$.

exhibited significant positive correlations between CV lag and sonority difference. The coronal C stimuli with differences in C voicing did not show the expected pattern. Also, the bilabial C stimuli with differences in vowel height did not consistently show the expected pattern. The non-significant result is potentially due to the same V measure for different vowel heights, as well as the imprecise C voicing coding.

The results for the Mandarin experiment can be found in Table 6 and Table 7. As in Table 6, most groups in Mandarin EMA data exhibited a significant positive correlation between CV lag and sonority difference. Pairwise comparison in Table 7 showed that the stimuli with bilabial C generally had significant positive correlations when the stimuli differed in voicing, nasality, and vowel height. In those cases, the stimuli had controlled nasality, voicing, as well as C place and manner.

4. Discussion. The study showed that there is a significant positive correlation between CV lag and sonority difference for both English monosyllabic words and Mandarin tone 4 words. The positive correlation was found when there is the same C and different V, as well as when there is the same V and different C. Furthermore, the claim was also supported by more controlled comparisons of labial and coronal consonants, as well as vowels of different heights.

The findings of the study can provide a basis for several phonological universals if we add a premise that humans prefer larger gestural lags in articulation. Firstly, the Sonority Dispersion Principle (SDP) states that in a syllable CV, the onset and nucleus differ from each other in sonority as much as possible (Clements 1990; Parker 2011). In other words, this principle requires that a CV syllable should have larger *sonority difference*. The SDP is potentially derivable from the finding of the current study along with another premise that a larger gestural lag is preferred, potentially for reasons of perceptual recoverability (Chitoran et al. 2002). For instance, it may be that a larger lag is correlated with more perceptual salience so it is preferred. It is also possible that a larger lag is preferred because it is easier to articulate. Ohala (1990) argued that some sequences may be disfavored due to their being difficult to articulate — it is possible that having a shorter gestural lag for gestures serves as the physical manifestation of articulatory difficulty to implement a certain ordered sequence. More work needs to be done to

	Pairwise comparison	Stimulus pair	Estimate (sonority difference)
Bilabial C	Nasality differ	mi4, bi4	10.18 ***
		mei4, bei4	17.62 ***
		ma4, ba4	17.93 ***
		man4, ban4	20.54 ***
	Voicing differ	bi4, pi4	10.57 ***
		bei4, pei4	6.95 *
		ba4, pa4	9.69 *
		ban4, pan4	8.11 *
	Vowel height differ	mi4, ma4	2.58
		bi4, ba4	16.72 **
		pi4, pa4	14.09 *
Coronal C	Nasality differ	nu4, du4	30.89 ***
		na4, da4	43.65 ***
	Voicing differ	tu4, du4	-3.31
		ta4, da4	1.12
	Vowel height differ	tu4, ta4	7.23
		su4, sa4	-8.47
		du4, da4	0.51
		nu4, na4	-18.67 **
		lu4, la4	-3.7

Table 7. Summary of EMA Mandarin pairwise comparisons. *** means that $p \leqslant 0.001$; ** means that $p \leqslant 0.01$; * means that $p \leqslant 0.05$.

assess articulatory ease, which has been elusive to define or study (Shariatmadari 2006).

Secondly, the findings of the study could be used to explain the SSP if we generalize the link to CC onset clusters. The SSP requires that a sonority rise (such as [pl]) is preferred in onsets over a sonority plateau (such as [pt]) which in turn is preferred over a sonority fall (such as [lp]) cross-linguistically (Sievers 1881, 1901; Greenberg 1965; Pike 1972; Hooper & Bybee 1976; Steriade 1982; Selkirk 1984; Clements 1990; Kenstowicz 1994; Blevins 1995; Parker 2002, 2011). If one generalizes the finding on CV sequences to CC sequences, one would predict that sonority rise has a larger lag than sonority plateau, which has a larger lag than sonority fall. This prediction has already been supported by Georgian (Crouch 2022; Crouch et al. 2023). If the premise that humans prefer larger gestural lag within a syllable is true, we would predict a phonological constraint that sonority rise is preferred over plateau over fall.

Thirdly, a related question would be how the current finding can inform the explanation of the Syllable Contact Law, which specifies that the structure A.B would be more preferable if a-b (a and b refer to the sonority of A and B respectively) is larger (Hooper & Bybee 1976; Murray & Vennemann 1983). We would argue that a general preference toward larger gestural lag within the syllable could derive the Syllable Contact Law. Consider the syllable $CV_1A.BV_2$, where there are two syllables CV_1A and BV_2, with A.B at the syllable boundary. If both syllables need to satisfy the requirement that larger lags are preferred in a syllable, then it comes as a consequence that a-b is larger. Since in each syllable, every segment sequence should satisfy the large lag requirement, V_2-b and a-V_1 should have larger sonority differences. This means that, on the sonority index, V_2 and a should be larger, and b and V_1 should be smaller. If a is large and b is small, and a-b would be larger.

Furthermore, the premises used above can also be used to provide a basis to explain why cross-linguistically, CV syllables are much more common than VC syllables (Ohala 1990; Tabain et al. 2004; Nam et al. 2009). Specifically, VC will have a shorter gestural lag than CV, which is disfavored if the premise that human prefers larger lag is true.

Moreover, the finding of the study allows us to partially understand why lower vowels are acoustically longer and have higher intensity than higher vowels (Lehiste 1970; Gordon et al. 2012). Following the main claim of the paper, since lower vowels are more sonorous, they will have a larger lag with the preceding consonant. Therefore, there is less overlap with the preceding consonant, and thus there is less acoustic "hiding" of the vowel. Consequently, a low vowel is likely to be acoustically longer and louder than a high vowel.

Additionally, our proposal leads to the prediction that lower vowels are perceptually and acoustically longer, making them better tone and stress holders than higher vowels. Such cases can be found in many languages. For example, Zuraw (2003) found that in Palauan, more sonorous vowels are dispreferred in unstressed syllables. Similarly, Gordon et al. (2012) found that in Armenian, Javanese, and Kwak'wala, the reduced phonological sonority of schwa relative to peripheral vowels is manifested in the rejection of stress on schwa. This indicates a positive correlation between sonority and stress — reduced sonority correlates to the absence of stress. Moreover, avoidance of stressed high vowels has been observed in Takia (Ross 2002, 2003; De Lacy 2007) — while the final syllable is stressed by default, if the final vowel is a high vowel, stress falls on a non-high vowel elsewhere in the word.

The relationship between sonority difference and lag could be one of the reasons why stress favors high sonority vowels. As observed by Gu (2023), Katsika (2016) and Katsika (2012), stressed vowels are correlated with larger gestural lags than unstressed ones. Since given the same consonant, vowels high in sonority are correlated with larger gestural lags than vowels low in sonority, this in turn leads to an acoustically and perceptually longer lower vowel, which therefore provides a better holder for stress.

Lastly, this current study has the potential to support a new speech production model that assumes sonority determines gestural coordination patterns. The assumption of this sonority-based speech production model is that all gestures coordinate according to the sonority differences with the gestures of the adjacent segment within a syllable. Consequently, the findings of the current study should be generalized to CC, CV, and VC sequences in syllables in all languages, where a positive correlation between sonority difference and gestural lag is predicted for cases beyond CV syllables.

References

Albert, Aviad. 2023. *A model of sonority based on pitch intelligibility*. BoD–Books on Demand.

Blevins, Juliette. 1995. The syllable in phonological theory. In *Handbook of phonological theory*, 206–244. Blackwell.

Chitoran, Ioana, Louis Goldstein & Dani Byrd. 2002. Gestural overlap and recoverability: Articulatory evidence from georgian. *Laboratory phonology* 7(4-1). 419–447.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology: Volume 1, Between the Grammar and Physics of Speech* 1. 283.

Clements, George N. 2005. Does sonority have a phonetic basis? comments on the chapter by vaux. 14 pp. In *In raimy, e. & cairns, c.(eds.), contemporary views on architecture and representations in phonological theory*, Citeseer.

Crouch, Caroline. 2022. *Postcards from the syllable edge: sonority and articulatory timing in complex onsets in georgian*: UC Santa Barbara dissertation.

Crouch, Caroline, Argyro Katsika & Ioana Chitoran. 2023. Sonority sequencing and its relationship to articulatory timing in georgian. *Journal of the International Phonetic Association* 1–24.

De Lacy, Paul. 2007. The interaction of tone, sonority, and prosodic structure. *The Cambridge handbook of phonology* 281. 281–308.

Durvasula, Karthik & Yichen Wang. 2023. Revisiting cv timing with a new technique to identify inter-gestural proportional timing. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* 2284–2288.

Gao, Man. 2008. *Mandarin tones: An articulatory phonology account*: Yale University dissertation.

Gordon, Matthew, Edita Ghushchyan, Bradley McDonnell, Daisy Rosenblum & Patricia A Shaw. 2012. Sonority and central vowels: A cross-linguistic phonetic study. *The sonority controversy* 219–256.

Greenberg, Joseph H. 1965. Some generalizations concerning initial and final consonant sequences. *Linguistics* 3(18). 5–34.

Gu, Yunting. 2023. Exploring the effect of stress on gestural coordination. *Proceedings of the Linguistic Society of America* 8(1). 5539.

Hooper, Joan B & Joan L Bybee. 1976. *An introduction to natural generative phonology*. Academic Press.

Kang, Yoongjung, Marc van Oostendorp, Colin J Ewen, Elizabeth Hume & Keren Rice. 2011. *The blackwell companion to phonology*.

Katsika, Argyro. 2012. *Coordination of prosodic gestures at boundaries in greek*. Yale University.

Katsika, Argyro. 2016. The role of prominence in determining the scope of boundary-related lengthening in greek. *Journal of phonetics* 55. 149–181.

Kenstowicz, Michael J. 1994. *Phonology in generative grammar*, vol. 7. Blackwell Cambridge, MA.

Lehiste, Ilse. 1970. Suprasegmentals, cambridge, massachusetts & london, uk.

Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford University Press.

Murray, Robert W & Theo Vennemann. 1983. Sound change and syllable structure in germanic phonology. *Language* 514–528.

Nam, Hosung, Louis Goldstein & Elliot Saltzman. 2009. Self-organization of syllable structure: A coupled oscillator model. *Approaches to phonological complexity* 16. 299–328.

Ohala, John J. 1990. Alternatives to the sonority heirarchy. In *Papers from the 26th regional meeting of the chicago linguistics society*, vol. 2, .

Parker, Steve. 2002. *Quantifying the sonority hierarchy*: University of Massachusetts at Amherst dissertation.

Parker, Steve. 2008. Sound level protrusions as physical correlates of sonority. *Journal of phonetics* 36(1). 55–90.

Parker, Steve. 2011. Sonority. *The Blackwell companion to phonology* 1–25.

Pike, Kenneth L. 1972. Phonetics: A critical analysis of phonetic theory and a technic for the practical description of sounds. .

R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Ross, Malcolm. 2002. Takia. the oceanic languages ed. by john lynch, malcolm ross & terry crowley, 216-248.

Ross, Malcolm. 2003. Seminar on takia, a papuanised austronesian language of papua new guinea.

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. *Language sound structure* .

Shariatmadari, David. 2006. Sounds difficult? why phonological theory needs ‘‘ease of articulation‘‘. *School of Oriental and African Studies Working Papers in Linguistics* 14. 207–226.

Shaw, Jason & Wei-rong Chen. 2019. Spatially conditioned speech timing: evidence and implications. *Frontiers in Psychology* 10. 2726.

Sievers, E. 1881. Grundzüge der phonetik: Breitkopf und hartel.

Sievers, Eduard. 1901. *Grundzüge der phonetik: zur einführung in das studium der lautlehre der indogermanischen sprachen*, vol. 1. Breitkopf & Härtel.

Smolensky, Paul. 1995. On the structure of the constraint component con of ug. *Handout of the talk presented at UCLA* .

Steriade, Donca. 1982. *Greek prosodies and the nature of syllabification*: Massachusetts Institute of Technology dissertation.

Tabain, Marija, Gavan Breen & Andrew Butcher. 2004. Vc vs. cv syllables: a comparison of aboriginal languages with english. *Journal of the International Phonetic Association* 34(2). 175–200.

Tiede, Mark. 2005. Mview: software for visualization and analysis of concurrently recorded movement data. *New Haven, CT: Haskins Laboratories* .

Zhang, Muye, Christopher Geissler & JA Shaw. 2019. Gestural representations of tone in mandarin: evidence from timing alternations. In *Proceedings of the 19th international congress of phonetic sciences, melbourne, australia 2019*, 1803–1807. Australasian Speech Science and Technology Association Inc Canberra, ACT.

Zuraw, Kie. 2003. Vowel reduction in palauan reduplicants. In *Proceedings of the 8th annual meeting of the austronesian formal linguistics association [afla 8]*, 385–398.