〈 Brief Report 〉

Compare Japanese-English bilinguals with Mandarin-English bilinguals: Focus on the production of English stop-stop sequences across word boundaries

Sha Liu

Kaye Takeda

Abstract:

Recent years have witnessed a growing interest in speech production by bilinguals with two first languages. The present study focuses on the production of stop-stop sequences across word boundaries by Japanese-English bilinguals and Mandarin-English bilinguals. Data from monolingual English speakers and from ESL speakers of Japanese and Mandarin were also collected for comparison. Although Japanese and Mandarin share a similar constraint on consonant sequences, statistically significant differences emerged from the two bilingual groups. In addition, differences between the monolingual group and the two bilingual groups are larger than expected. Results here appear to show that bilinguals may have their unique features and be different from monolinguals in both languages.

Key words: bilingual, stop-stop sequence, ESL, monolingual, transfer, statistical analysis

1. Introduction

Cross-linguistic interaction has been used to explain differences between monolinguals and bilinguals (Dijkstra et al. 1998; Van Hell and Dijkstra 2002; Von Studnitz and Green 2002; Wu and Thierry 2010a, 2010b; Kroll et al. 2012; Simonet 2014; see also the review in Hambly et al. 2013). The literature claims that transfer exists between the two language systems and structures in bilinguals: even if only one language is utilized, both languages are constantly and jointly active to some degree. However, even scholars who claim that transfer between the two language systems exists in bilinguals cannot agree with each other about how phonological systems interact and what factors influence interactions, e.g., the direction of transfer and the extent of transfer (see e.g., Johnson 1989; Baker 2001; Paradis 2004; Blumenfeld and Marian 2007; Fabiano-Smith and Barlow 2010). These inconsistencies in previous studies spurred us to focus on the production of stop-stop sequences across word boundaries by Japanese-English bilinguals in Liu and Takeda (2024). Our results showed that the difference in terms of duration ratio between the monolingual and Japanese-English bilingual groups has been rejected as statistically significant ($p > 0.99$). However, the Japanese-English bilingual group does not share any common factors affecting duration ratios with the monolingual group. This suggests that the differences between monolinguals and bilinguals are more complex than previously assumed.

To further examine our conclusion drawn from Japanese-English bilinguals, we enrolled another group of bilinguals in the present study, Mandarin-English bilinguals. Similar to Japanese, Mandarin generally does not allow consonant sequences either. The most complex syllable structure in Mandarin is CGVX, where *G* refers to a glide and *X* either a nasal or the second part of a long vowel or a diphthong (Wang and Chang 2001; Duanmu 2011, 2016; Triskova 2011; Zhao and Berent 2016). Mandarin only exhibits consonant clusters across word boundaries beginning with a nasal consonant. Previous studies have pointed out that Mandarin speakers have substantial difficulty in pronouncing English consonant sequences (Chen and Chung 2008). For example, Weinberger (1987) reports the negative relationship between the size of consonant sequences and the accuracy of English consonant sequence production by native Mandarin speakers. Anderson (1987) more specifically states that Mandarin speakers have difficulty with English consonant sequence production and prefer to delete consonants in the coda position. Broselow et al. (1998) claim that Mandarin speakers favor insertion for English consonant sequences in monosyllabic words and deletion in disyllabic words. Wang (1995) focuses on English coda stop consonants /p, t, k, b, d, g/ and reports that Mandarin speakers have two major problems in their speech production: (i) the epenthesis of a vowel after a coda stop; and (ii) the deletion of a coda stop.

In this paper, we take Mandarin-English bilinguals (hereafter the ME group) as our focus and compare their production of stop-stop sequences across word boundaries with that of monolinguals of English (henceforth the monolingual group) and of native Mandarin speakers who speak English as a second language (hereafter the MESL group). If cross-linguistic interaction is the reason for the differences between monolinguals and bilinguals as reviewed at the beginning of this section, the transfer from the language pattern of Mandarin to the production in English is expected to be observed. More specifically, the ME group is expected to present intermediate results between the monolingual and MESL groups in terms of all or at least most of measures of interest here. We examine the same stop-stop sequences across word boundaries as Liu and Takeda (2024) do to have an easy comparison with their results from Japanese-English bilinguals (henceforth the JE group). The heterorganic sequences under analysis include [d#b], [d#g], [d#p], [d#t], [k#p], [k#t], [p#t], [p#k], [t#d], [t#k], and [t#p]; the homorganic sequences are /d#d/, /d#t/, /k#k/, /t#d/, and /t#t/. In the following, we term stop-stop sequences across word boundaries as stop-stop sequences for ease of exposition.

2. Methods

Concerning the monolingual group, we used the recordings and data in Liu and Takeda (2024) and thus did not enroll any new monolingual English speakers. The three monolingual native speakers of English were born and brought up in California. They were also residents in California at the time of recording. They could not make effective speech in any language other than English.

For the present study, we enrolled three ME bilinguals and three MESL speakers. The three ME bilinguals were all born in China and moved to California as infants. They acquired Mandarin and

English naturalistically and are fluent in both Mandarin and English. They were also residents in California at the time of recording. We limited speakers in the ME group to Californian English speakers to reduce the potential influences of different English accents as much as possible. The three MESL speakers were all born and brought up in the central part of China. They have not learned English in their early childhood, received English language education at school, and have not lived in any English-speaking country for more than one month. Their accents are not markedly different from standard Mandarin. The aim is also to reduce the possible influences of different Mandarin accents.

The monolinguals and the ME bilinguals were around 30 to 35 years old at the time of recording and college graduates. All MESL speakers were college students and were just over 20 years old at the time of recording. Although the MESL speakers are not of the same age as the monolinguals and EM bilinguals, the age difference is not remarkably large. The gender breakdown in each group is the same: one male and two females.

Similar to the experiment in Liu and Takeda (2024), all speakers were given the text from the PAC project, *Christmas Interview of a Television Evangelist*, long before their recording to get familiar with it. They were instructed to practice the passage in their normal voice and at a rate that they felt natural and comfortable until they could read the passage fluently. Before the recording, they were also instructed that they should repeat the whole sentence if they made a mistake. Recordings were made on the second author's iPhone 6 in quiet rooms and later converted to wav format for acoustic analysis on Praat. The first author segmented and labelled the recorded speeches. This procedure was carried out on speech waveforms and wideband spectrograms generated on Praat. Pauses between intonation phrases were excluded from the analysis. The segmental boundaries were identified generally by taking spectral transitions into consideration. A total of 373 sound files were acoustically analyzed, with the specific breakdown as follows: 162 tokens from the monolingual group, 124 tokens from the ME group, and 87 tokens from the MESL group.

3. Results

In this section, we compare results from the monolingual, ME, and MESL groups in terms of stop-stop sequence release and duration ratio respectively.

3.1. Stop-stop sequence release

Stop-stop sequence release refers to the release of the first consonant in a sequence under discussion, e.g., the release of /d/ in the sequence /d#b/. As reported in previous studies, factors related to the release of stop-stop sequences across word boundaries mainly include place order effect, word frequency, the relative frequency of the first and second words concerned, and stress (see e.g., Hardcastle and Roach 1979; Zsiga 2000; Walker 2008). Concerning the factor of place order effect, the possibility of the release of the first stop in a stop-stop sequence increases from homorganic sequences, front-to-back sequences, to back-to-front sequences according to Catford (1977; see also Browman and Goldstein 1990; Byrd 1992; Ladefoged 1993; Zsiga 2000). Concerning the frequency factor, first-word frequency and second-word frequency were considered since the frequency factor is disputable in speech production (Hooper 1976; Aylett and Turk 2004; Gahl 2008; Gahl et al. 2012; Coetzee and Kawahara 2013). The raw data of frequency were not normally distributed as graphically shown by the QQ plot generated on the GraphPad Prism version 8.0.0 for Windows (GraphPad Software, San Diego, CA; hereafter GraphPad software). Hence normalization was carried out on raw frequency data by use of the GraphPad software. The two factors were termed *normalized first-word frequency* and *normalized second-word frequency*. If the frequency has a role in speech production, it appears plausible that the relative frequency of the two words concerned also has a role. Thus, the factor *relative frequency of the first and second words* (henceforth *relative frequency*) was configured with two levels: *low/high* means that the first word of a stop-stop sequence has a lower frequency than the second word; *high/low* refers to a reversed pattern. Statistical analysis results in terms of stop-stop sequence release for each group are reported in Table 1.

Table 1 The binary logistic regression results in terms of stop-stop sequence release from the monolingual, ME, and MESL groups

	Factor level	Estimate	Std. Error	$\|z\|$	*p*	
The monolingual group						
Intercept		4.59	1.06	4.32	<0.00	***
Place order effect	Homorganic	0.47	0.08	0.73	0.04	*
	Front/Back					
	Back/Front					
Normalized first-word frequency	Numerical data	0.0001	0.00	2.24	0.03	*
Normalized second-word frequency	Numerical data	-0.0002	0.00	4.26	<0.00	***
Relative frequency	Low/High	-3.88	0.93	4.16	0.12	
	High/Low					
Stress	Unstressed-stressed	-3.59	0.86	4.18	<0.00	***
	Stressed-unstressed					
The ME group						
Intercept		2.66	1.42	1.87	0.03	*
Place order effect	Homorganic	3.71	1.27	2.93	0.00	**
	Front/Back					
	Back/Front					
Normalized first-word frequency	Numerical data					
Normalized second-word frequency	Numerical data	-0.00	0.00	1.68	0.09	
Relative frequency	Low/High	-6.49	27.40	2.39	0.02	*
	High/Low					
Stress	Unstressed-stressed	0.81	0.62	1.31	0.19	
	Stressed-unstressed					
The MESL group						
Intercept		-2.09	1.01	2.06	0.04	*
Place order effect	Homorganic	-.0.48	1.89	3.74	0.43	
	Front/Back					
	Back/Front					
Normalized first-word frequency	Numerical data	0.0002	0.00	1.63	0.05	*
Normalized second-word frequency	Numerical data	-0.0001	0.00	3.59	0.04	*
Relative frequency	Low/High	-0.12	0.98	2.83	0.75	
	High/Low					
Stress	Unstressed-stressed	2.49	0.79	3.14	0.00	***
	Stressed-unstressed					

Notes: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.
$|t|$ stands for the absolute value of *t* as given in the GraphPad software.

The dependent variable for the binary logistic regression model was stop-stop sequence release. Each sequence was coded for the presence or absence of a release burst. If the burst was visible on the spectrogram, it was tagged as *released*. Otherwise, it was tagged as *unreleased*. The statistical analysis results in Table 1 show that the factors of *place order effect*, *normalized first-word frequency*, *normalized second-word frequency*, and *stress* are significantly related to stop-stop sequence release in the monolingual group ($p = 0.04$, $p = 0.03$, $p < 0.00$, $p < 0.00$). The factors of *place order effect* and *relative frequency* are significantly related to stop-stop sequence release in the ME group ($p = 0.00$, 0.02). A stop-stop sequence is more likely to be released from the homorganic, front/back, to back/front level in the ME group, as indicated by its positive coefficient (Estimate = 3.71). Although the monolingual group has a similar tendency as the ME group in terms of place order effect, this factor has a much smaller effect in the monolingual group than in the ME group (Estimate = 0.47, 3.71). The negative value of the coefficient of the relative frequency factor shows that a stop-stop sequence is more likely to be released in the ME group when the second word in a sequence has a higher frequency than the first word (Estimate = -6.49).

For the MESL group, *normalized first-word frequency*, *normalized second-word frequency*, and *stress* have emerged as statistically significant factors ($p = 0.05$, 0.04, 0.00). The positive and negative values of the coefficients for *normalized first-word frequency* and *normalized second-word frequency* respectively show that the higher the frequency of the first word is and the lower the frequency of the second word is, the more possible for a related stop to be released, although the relationship is relatively weak as indicated by the absolute values of their respective coefficients (Estimate = 0.0002, -0.0001). The relationship between normalized first-word frequency and normalized second-word frequency and stop-stop sequence release is generally the same between the MESL group and the monolingual group. From here on, we will omit discussion related to factors whose coefficients are close to zero since the effects of these factors are negligible (Anderson 2014, Baer 2019). The stress factor did not emerge as statistically significant in the bilingual group ($p = 0.19$), while it did in the monolingual and MESL groups ($p < 0.00$, $p = 0.00$). The relationship between stress and stop-stop sequence release is reversed between the monolingual and MESL groups (Estimate = -3.59, 2.49).

A comparison of results from the three groups in Table 1 shows the following points: (1) the ME group shares only one common factor, place order effect, with the monolingual group; (2) the ME group does not share any statistically significant factors with the MESL group; (3) although stress emerged as statistically significant in both the monolingual and MESL groups, this factor has reversed effects in the two groups. The chi-square test performed using the GraphPad software for stop-stop sequence release shows that the differences between groups are statistically significant ($p < 0.00$). The results between every two groups with the adjusted p-values for Bonferroni correction show that the monolingual and MESL groups have a statistically significant difference ($p = 0.00$), and so do the ME and MESL groups ($p < 0.00$). The difference between the monolingual and ME groups is not statistically significant ($p > 0.99$). The preliminary results here appear to show that the ME group resemble more the monolingual group than the MESL group.

3.2. Duration ratio

We follow Zsiga (2000: 70) and calculate *duration ratio* by "a comparison of the duration of the cluster with the duration of the two consonants (in corresponding word-final and word-initial position) occurring singly between vowels." The multiple linear regression results in terms of duration ratio for the monolingual, ME, and MESL groups are given in Table 2.

Table 2 The multiple linear regression results in terms of duration ratio from the monolingual, ME, and MESL groups

	Factor level	Estimate	Std. Error	|t|	*p*	
The monolingual group						
Intercept		1.05	0.12	8.59	0.00	***
Place order effect	Homorganic Back/Front Front/Back	0.00	0.02	0.10	0.92	
Normalized first-word frequency	Numerical data	-0.0003	0.00	5.71	0.00	***
Normalized second-word frequency	Numerical data	-0.0004	0.00	1.94	0.14	
Relative frequency	Low/High High/Low	0.31	0.15	2.13	0.03	*
Stress	Unstressed-stressed Stressed-unstressed	1.26	0.21	1.28	0.20	
The ME group						
Intercept		1.38	0.09	16.00	<0.00	***
Place order effect	Homorganic Back/Front Front/Back	0.18	0.10	1.81	0.08	
Normalized first-word frequency	Numerical data	-0.0001	0.00	3.37	0.00	***
Normalized second-word frequency	Numerical data	-0.0004	0.00	7.50	<0.00	***
Relative frequency	Low/High High/Low	-4.96	1.26	3.93	0.00	***
Stress	Unstressed-stressed Stressed-unstressed	0.39	0.07	5.65	<0.00	***
The MESL group						
Intercept		1.43	0.01	11.10	<0.00	***
Place order effect	Homorganic Back/Front Front/Back	-0.03	0.01	2.23	0.03	*
Normalized first-word frequency	Numerical data					
Normalized second-word frequency	Numerical data	-0.0002	0.00	52.50	<0.00	***
Relative frequency	Low/High High/Low	-2.68	0.16	16.66	<0.00	***
Stress	Unstressed-stressed Stressed-unstressed	0.0002	0.01	0.05	0.96	

Notes: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.
$|t|$ stands for the absolute value of t as given in the GraphPad software.

Independent factors for this model are generally the same as those in Table 1. The only exception is *place order effect*. Duration ratio increases from homorganic sequences, back-to-front sequences, to front-to-back sequences according to Zsiga (2003). Thus, *place order effect* is configured with three levels, *homorganic*, *back/front*, and *front/back*. The order of the three levels here is different from that in Table 1 for the statistical analysis of stop-stop sequence release.

The factors of *normalized first-word frequency* and *relative frequency* have emerged as significant for the monolingual group (p = 0.00, 0.03). However, since the coefficient of *normalized first-word frequency* is close to zero, we will omit discussion related to it (Estimate = -0.0003). The positive value of the coefficient of the relative frequency factor means that the duration ratio of a stop-stop sequence tends to be larger when the first word involved in the sequence has a higher frequency (Estimate = 0.31).

The factors of *normalized first-word frequency*, *normalized second-word frequency*, *relative frequency*, and *stress* emerged as significant for the ME group (p = 0.00, p < 0.00, p = 0.00, p < 0.00). The negative coefficient value of *relative frequency* presents the following pattern: duration ratio tends to be larger when the second word has a higher frequency than the first word (Estimate = -4.96). The positive coefficient value of *stress* states that duration ratio tends to be larger when the stop-stop sequence has a stressed-unstressed pattern (Estimate = 0.39).

The model output shows that *place order effect*, *normalized second-word frequency*, and *relative frequency* emerged as statistically significant for the MESL group. *Place order effect* yields a negative effect on duration ratio from the homorganic, back/front, to front/back level (Estimate = -0.03). *Relative frequency* has a negative relationship with duration ratio: duration ratio tends to be larger when the second word has a higher frequency than the first word (Estimate = -2.68).

Although relative frequency emerged as statistically significant both in the monolingual and ME groups, the coefficients of this factor in the two groups indicate different tendencies: duration ratio tends to decrease from the low/high to high/low level in the ME group (Estimate = -4.96), while duration ratio has a reversed tendency in the monolingual group (Estimate = 0.31). In contrast, the ME and MESL groups not only share the common statistically significant factor of relative frequency, but also similar coefficients (Estimate = -4.96, -2.68). The one-way ANOVA test shows statistically significant differences between the monolingual, ME, and MESL groups in terms of duration ratio (p = 0.01). Post hoc comparisons using Dunnett's T3 multiple comparisons test show statistically significant differences between the monolingual and ME groups (p = 0.04), between the monolingual and MESL groups (p = 0.01), and between the ME and MESL groups (p = 0.02). Results in terms of duration ratio appear to show larger differences between the ME and monolingual groups than results in terms of stop-stop sequence release.

3.3. Individual variation

Intra-group differences in terms of stop-stop sequence release were examined using the chi-square test. The three speakers in the monolingual group had a statistically significant difference (p = 0.00). The adjusted p-values for Bonferroni correction show statistically significant differences between two comparisons (p = 0.00, 0.02), but not between the last comparison (p > 0.99). There were no statistically significant differences between every two participants in the ME group (p = 0.20). The three speakers in the MESL group show statistically significant differences in terms of stop-stop sequence release (p = 0.00). The adjusted p-values for Bonferroni correction show statistically significant differences between two comparisons ($p = 0.02$, $p < 0.00$), but not between the last comparison ($p = 0.32$).

In terms of duration ratio, a one-way ANOVA test showed that the three speakers in the monolingual group had a statistically significant difference ($p = 0.03$). Post hoc Holm-Sidak's multiple comparisons tests showed statistically significant differences in two comparisons ($p = 0.05$, 0.05), but not in the last comparison ($p = 0.78$). The three speakers in the ME group had statistically significant differences ($p = 0.00$), but one comparison did not have statistically significant differences ($p = 0.27$). The other two comparisons had statistically significant results ($p = 0.02$, $p = 0.00$). The three speakers in the MESL group did not show a statistically significant result ($p = 0.09$). The results here show that intra-group differences exist although general tendencies are the same in each group.

4. Discussion and conclusion

Previous studies have stated that consonant sequence production poses problems for native Japanese speakers and native Mandarin speakers. Strategies that they may use include vowel insertion and consonant deletion. However, the present study has very limited data on vowel insertion and consonant deletion. One possible explanation is that we asked participants to practice before recording and read at a rate that they were comfortable. Since participants could prepare beforehand and had certain freedom in speech rate, they were able to avoid mistakes to a large extent.

Stop-stop sequence release has not presented statistically significant differences between the monolingual and ME groups ($p > 0.99$), while duration ratio has ($p = 0.04$). Although the results from the ME group appear to be closer to those from the monolingual group, they are not as close as expected. For example, in terms of stop-stop sequence release, although the ME group shares the common statistically significant factor of place order effect with the monolingual group, the effect of this factor is much stronger in the ME group than in the monolingual group (Estimate = 3.71, 0.47). In terms of duration ratio, the only common statistically significant factor of relative frequency between the ME and monolingual groups has reversed effects in the two groups (Estimate = -4.96, 0.31).

Liu and Takeda (2024) compare the production of stop-stop sequences across word boundaries by the same monolingual English group as in this study, a Japanese-English bilingual group (hereafter the JE group), and native Japanese speakers who speak English as a second language (henceforth the JESL group). The differences between the monolingual and JE groups were as follows: (i) statistically significant differences in terms of stop-stop sequence release exist between the monolingual and JE groups ($p < 0.00$); (ii) the difference in terms of duration ratio between the monolingual and JE groups has been rejected as statistically significant ($p > 0.99$). Concerning the ME and monolingual groups, a statistically significant difference emerged in terms of duration ratio ($p = 0.04$). Our results indicate that differences exist between monolinguals and bilinguals, but the differences are not extremely large. However, if we take a closer look at factors related to stop-stop sequence production, we will notice larger differences between monolinguals and bilinguals. To exemplify, the ME group only shares one common factor with the monolingual group in terms of stop-stop sequence release.

Concerning a comparison between the JE group and the ME group, the chi-square test performed on the GraphPad software shows that the two groups have statistically significant differences in terms of stop-stop sequence release ($p < 0.00$). The unpaired t-test carried out on the same software shows that statistically significant differences also emerged in terms of duration ratio ($p < 0.00$). The two groups do not share many common statistically significant factors either. In terms of stop-stop sequence release, the two groups only share one common factor, *relative frequency*. However, even this one common factor has reversed effects in the two groups (Estimate = 0.29, -6.49). In terms of duration ratio, the two groups have one common factor, *normalized second-word frequency* and similar coefficients (Estimate = -0.00, -0.00). However, since the coefficients are close to zero, this common factor can be overlooked.

The maximal syllable structure is CGVX in Mandarin and C(j)VX in Japanese. Moreover, speakers in both groups spent their formative years in California and are college graduates. Both groups are composed of two females and one male of a similar age. It seems plausible that results from the ME group should be similar to those from the JE group following the claim of cross-linguistic interaction. Our results have demonstrated that it does not follow to attribute differences between the ME and JE groups solely to the differences between Mandarin and Japanese: cross-linguistic interaction cannot explain every difference between monolinguals and bilinguals. It appears that certain aspects of bilinguals' phonetic performance may be different from that of monolinguals in both languages. Factors involved may include suprasegmental pattern, i.e. stress, phonotactic constraint, frequency, and so on. An intertwining of all these factors may have presented a monolingual-like pattern with respective unique features in bilinguals.

The focus of this paper is the production of stop-stop sequences across word boundaries. In terms of this, English is quite distinct from Japanese and Mandarin. This may be part of the reasons that bilinguals in this paper have statistically significant differences from English monolinguals. If the focus has been given to a point that does not have such a sharp contrast between English, Japanese, and Mandarin, results may not be as distinctive as they are in this paper. Further research into this aspect will be left for future study. We understand that this study is based on small group sizes. More research with larger groups of bilinguals is needed to confirm the current findings.

***Acknowledgements**

For help in getting this article to its final form, our special gratitude goes to Prof. Jacques Durand for his advice on acoustic analysis, to Prof. Eiji Yamada and Prof. Hajime Takeyasu for advice and discussion, to Prof. Greg Bevan and Prof. Robert Long for editing our paper, and to the editors of the present journal for detailed and helpful feedback. All remaining errors are our responsibility. This work was funded by JSPS Grant-in-Aid for Early-Career Scientists (KAKENHI-PROJECT-20K13072).

References

Anderson, Alan. 2014. Business statistics for dummies. New Jersey: John Wiley & Sons.

Anderson, Janet. 1987. The markedness differential hypothesis and syllable structure difficulty. In *Interlanguage Phonology: The Acquisition of a Second Language Sound System*, ed. by Georgette Ioup and Steven Weinberger, 279-291. Cambridge, MA: Newbury House.

Aylett, Matthew and Turk, Alice. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47(1): 31-56.

Baer, Tobias. 2019. Understand, manage, and prevent algorithmic bias: A guide for business users and data scientists. New York: Springer.

Baker, Colin. 2001. *Foundations of bilingual education and bilingualism.* Cambridge: Multilingual Matters LTD.

Blumenfeld, Henrike and Marian, Viorica. 2007. Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes* 22(5): 633-660.

Browman, Catherine P. and Goldstein, Louis M. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, ed. by J. Kingston and M. E. Beckman, 341-376. Cambridge: Cambridge University Press.

Byrd, Dani. 1992. *Perception of Assimilation in Consonant Clusters: A Gestural Model*. *Phonetica* 49: 1-24.

Catford, John. C. 1977. *Fundamental problems in phonetics*. Bloomington: Indiana University Press.

Chen Hsueh-Chu, and Raung-Fu Chung. 2008. Interlanguage analysis of phonetic timing patterns by Taiwanese learners. *Concentric: Studies in Linguistics* 34: 79-108.

Coetzee, Andries W. and Kawahara, Shigeto. 2013. Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 31: 47-89.

Dijkstra, T., van Jaarsveld, H., and ten Brinke, S. 1998. Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition* 1: 51-66.

Duanmu, San. 2011. Chinese syllable structure. In *The Blackwell Companion to Phonology*, Volume 5, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice, 2151-2777. Malden, MA: Wiley-Blackwell.

Duanmu, San. 2016. Syllable Structure. In *Encyclopedia of Chinese Language and Linguistics*, Volume 4, ed. by Rint Sybesma, 230-236. Leiden: Brill.

Fabiano-Smith, Leah and Barlow, Jessica A. 2010. Interaction in bilingual phonological acquisition:

Evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism* 13(1): 1-19.

Gahl, Susanne. 2008. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84(3): 474-496.

Gahl, Susanne, Yao, Yao and Johnson, Keith. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language* 66(4): 789-806.

GraphPad Prism. Version 8.0.0 for Windows. San Diego, California: GraphPad Software, 2019. Computer software.

Hambly, Helen, Wren, Yvonne, McLeod, Sharynne, and Roulstone, Sue. 2013. The influence of bilingualism on speech production: A systematic review. *International Journal of Language & Communication Disorders* 48(1): 1-24.

Hardcastle, William, and Roach, Peter. 1979. An instrumental investigation of coarticulation in stop consonant sequences. In *Current issues in the phonetic sciences 1*, ed. by Hollien, Harry and Hollien, Patricia, 531-540. Amsterdam: John Benjamins.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphological change. In *Current progress in historical linguistics*, ed. by William M. Christie, 95-105. Amsterdam: North-Holland.

Johnson, Janice. 1989. Factors related to cross-language transfer and metaphor interpretation of bilingual children. *Applied Psycholinguistics* 10(2): 157-177.

Kroll, Judith F., Dussias, Paola E., Bogulski, Cari A., and Kroff, Jorge R. Valdes. 2012. Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. *Psychology of learning and motivation* 56: 229-262.

Ladefoged, Peter. 1993. A Course in Phonetics, 3rd edition. New York: Harcourt, Brace, Jovanovich.

Liu, sha and Kaye Takeda. 2024. The production of stop-stop sequences across word boundaries: A preliminary report. *Kyushu Studies in English Literature* 40: 1-16.

Paradis, Michel. 2004. A Neurolinguistic Theory of Bilingualism. Amsterdam: John Benjamins.

Simonet, Miquel. 2014. Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics* 43: 26-37.

Triskova, Hana. 2011. The Structure of the Mandarin Syllable: Why, When and How to Teach It. *Oriental Archive* 79(1): 99-134.

Van Hell, J. G., and Dijkstra, A. 2002. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review* 9: 780-789.

Von Studnitz, Roswitha E., and Green, David. W. 2002. Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition* 5(1): 1-23.

Walker, James A. 2008. *Form, Function and Frequency in Phonology: (t/d)-deletion in Toronto*

English. Presented at New Ways of Analyzing Variation 37, Rice University, Houston, TX.

Wang, Chilin. 1995. The acquisition of English word-final obstruents by Chinese speakers. PhD dissertation, SUNY at Stony Brook, New York, USA.

Wang, H. Samuel and Chang, Chih-ling. 2001. On the status of the prenucleus glide in Mandarin Chinese. *Language and Linguistics* 2(2): 243-260.

Weinberger Daniel R. 1987. Implications of normal brain development for the pathogenesis of schizophrenia. *Archives of General Psychiatry* 44: 660-669.

Wu, Yanjing, and Thierry, Guillaume. 2010a. Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience* 30(22): 7646-7651.

Wu, Yanjing, and Thierry, Guillaume. 2010b. Investigating bilingual processing: the neglected role of language processing contexts. *Frontiers in Psychology* 1: 1-6. Accessed January 13. https://www.frontiersin.org/articles/10.3389/fpsyg.2010.00178/full.

Zhao, Xu and Iris Berent. 2016. Universal Restrictions on Syllable Structure: Evidence from Mandarin Chinese. *Journal of Psycholinguistic Research volume* 45: 795-811.

Zsiga, Elizabeth C. 2000. Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28: 69-102.

Zsiga, Elizabeth C. 2003. Articulatory Timing in a Second Language: Evidence from Russian and English. *Studies in Second Language Acquisition* 25: 399-432.