

*Biomechanically Preferred Consonant-Vowel Combinations Fail to Appear in Adult
Spoken Corpora*

D. H. Whalen^{1,2}, Sara Giulivi^{1,3}, Hosung Nam¹, Andrea G. Levitt^{1,4}, Pierre
Hallé^{1,5} and Louis M. Goldstein^{1,6}

¹Haskins Laboratories

²City University of New York

³University of Bologna

⁴Wellesley College

⁵CNRS/Sorbonne-Nouvelle

⁶University of Southern California

Corresponding Author:
D. H. Whalen
Haskins Laboratories
300 George St., Suite 900
New Haven, CT 06511
203-865-6163, ext. 234
whalen@haskins.yale.edu

Acknowledgments:

This work was supported by NIH grants DC-000403 and DC-002717 to Haskins Laboratories. We thank Aude Noiray, Stephen Frost, Barbara L. Davis and an anonymous reviewer for helpful comments.

Key words: phonotactics, word frequency, type/token, Articulatory Phonology, Frame then Content

Running head: CV combinations in adult corpora

Abstract

Certain consonant/vowel (CV) combinations are more frequent than would be expected from the individual C and V frequencies alone, both in babbling and, to a lesser extent, in adult language, based on dictionary counts: Labial consonants co-occur with central vowels more often than chance would dictate; coronals co-occur with front vowels, and velars with back vowels (Davis & MacNeilage, 1994). Plausible biomechanical explanations have been proposed, but it is also possible that infants are mirroring the frequency of the CVs that they hear. As noted, previous assessments of adult language were based on dictionaries; these "type" counts are incommensurate with the babbling measures, which are necessarily "token" counts. We analyzed the tokens in two spoken corpora for English, two for French and one for Mandarin. We found that the adult spoken CV preferences correlated with the type counts for Mandarin and French, not for English. Correlations between the adult spoken corpora and the babbling results had all three possible outcomes: significantly positive (French), uncorrelated (Mandarin), and significantly negative (English). There were no correlations of the dictionary data with the babbling results when we consider all nine combinations of consonants and vowels. The results indicate that spoken frequencies of CV combinations can differ from dictionary (type) counts and that the CV preferences apparent in babbling are biomechanically driven and can ignore the frequencies of CVs in the ambient spoken language.

1. Introduction

Although the sound systems of the world's languages show a great deal of diversity, they also exhibit some underlying commonalities. Comparisons of phonological inventories have shown that consonants at three places of articulation (labial, coronal and velar) are nearly universal, as are vowels at three degrees of fronting (front, central, back) (Jakobson, 1968; Lindblom, 1986; Maddieson, 1984, 1997). Such universal tendencies have been attributed variously to linguistic features (Jakobson, 1968), listener constraints (Lindblom, 1990), or coupled articulatory and auditory factors (Honda, 1996), but determining the source of these universal tendencies remains an active area of investigation.

Phonotactic combinations of consonants and vowels have also been found to show some universal tendencies. One surprising pattern was first observed by MacNeilage and Davis (Davis & MacNeilage, 1995; MacNeilage, 1998; MacNeilage & Davis, 1990) in infant babbling. When they plotted the incidence of each of the three consonant types mentioned above against the three vowel types, they found three preferred patterns of associations between adjacent consonants and vowels out of the nine possible combinations. The preference was evident when the ratio for the observed versus expected combinations was greater than 1.0. Labial consonants co-occurred with central vowels more often than would be expected by chance, as did coronal consonants with front vowels and dorsal consonants with back vowels. Infants explore a broad range of phonetic possibilities during babbling, so finding such regularities is unexpected, even though babblers do exhibit some universal tendencies as well, such as the use of point vowels [i a u] and pulmonic egressive stops (e.g., Oller, 2000; Vihman, Macken, Miller, Simmons, & Miller, 1985).

Because the CV combination patterns are observed before anything resembling true language is in evidence, biomechanical explanations seem plausible. MacNeilage and Davis proposed the ‘Frame then Content’ (F/C) account in which jaw movement plays the primary role (see also MacNeilage & Davis, 2011), while an account within Articulatory Phonology (Giulivi, Whalen, Goldstein, Nam, & Levitt, 2011; Whalen, Giulivi, Goldstein, Nam, & Levitt, 2011) proposes that synergy among articulators is the foundation.

Although a biomechanical explanation is possible for these CV preferences, there is a simpler explanation of their occurrence that has not yet been ruled out: Infants could be sensitive to the presence of such patterns in adult speech and therefore incorporate them into their babbling. The ambient language is known to influence some aspects of babbling even very early stages (Boysson-Bardies, Hallé, Sagart, & Durand, 1989; Whalen, Levitt, & Wang, 1991), so perhaps this pattern reflects only such an influence. Indeed, two studies have shown that the three preferred co-occurrence patterns are also somewhat preferred in adult languages. Maddieson and Precoda (1992) studied type frequencies (dictionary counts) in five languages with small phoneme inventories (Hawaiian, Rotokas, Pirahã, Eastern Kadazan and Shipibo). They did not find any direct evidence of preference for articulatory “convenience” (p. 55), that is, synergy; however, their raw numbers (“vowel deviation scores”) are consistent with the ratios found in babbling. MacNeilage et al. (2000) found that the CV syllables preferred in babbling were also preferred in dictionary counts of 10 languages, namely English, Estonian, French, German, Hebrew, Japanese, New Zealand Maori, Quichua, Spanish and Swahili. They found that the ratios of observed to expected frequencies are generally greater than 1.0 for the CV types that are favored in babbling and early speech: labial consonant-

central vowel (in 7 languages out of 10), coronal consonant-front vowel (in 7 languages out of 10) and dorsal consonant-back vowel (in 8 languages out of 10). Rousset (2004) found similar results for 15 languages: 11, 10 and 11 out of 15 for labial, coronal and velar CVs respectively (p. 194). Thus there is some reason to think that the adult language might be providing a template for the CV preference pattern in babbling.

The frequency with which an infant will hear certain syllables, however, is based on the frequency of words in the language (a "token" count) rather than on the total number of syllables within the words of a language (a "type" count). Some frequent words in English, for example, have preferred combinations ("but", "said", "go"), but many do not ("be", "do", "get"). Thus the frequency of the CVs that the infant hears could differ between dictionaries and spoken corpora. If so, then we need to compare the babbling ratios to those of the ambient spoken language.

Infants pay special attention to speech that is directed at them (Infant Direct Speech, or IDS) with modifications from typical adult-directed speech (ADS) (e.g., Fernald, 1985; Werker, Pegg, & McLeod, 1994), but they are exposed to more adult-directed speech than they are to infant-directed speech (van de Weijer, 2001) and can learn from it (Akhtar, Jipson, & Callanan, 2001; Oshima-Takane, 1988). In terms of segmental frequency, however, there appears to be little difference in which register is used. A recent large study of IDS in Korean and English found only a single, nonsignificant trend toward a segmental difference between the occurrence of high back vowels in English (S. A. S. Lee, Davis, & MacNeilage, 2010). There were no differences in segmental frequencies reported between Korean IDS and ADS. Speaking style also seems not to affect the frequency with which phonemes occur (Mines, Hanson, & Shoup, 1978). Therefore, it appears that the CV ratios in the

infant's language environment can be found in the analysis of adult spoken language corpora.

The one previous study of language corpora that we have found was less supportive of the role of preferred combinations in adult speech. Written data from five languages (Finish, Turkish, Latin, Latvian and Setswana) were studied by Janson (1986). He found support in each language for a dental/alveolar place for the C with front vowels and velars with back vowels, both in accord with the babbling results. Labials, however, also preferentially patterned with back vowels. Most of the corpora were relatively small (Latin: 11,803 CVs, Turkish: 33,135 CVs; Latvian: 16,665; Setswana: 20,325), while one was substantial (Finnish: 125,126 CVs). All sources were written texts as well, which could introduce differences from the spoken language. Janson's results must therefore be taken as an important first step but insufficient to provide a clear answer to the question of how CVs pattern in adult spoken languages.

Although the F/C account is meant only to describe the pattern of the preferred combinations (those on the diagonal in a 3x3 table of place of articulation for the consonant and the vowel), one further aspect of CV combinations is that there is systematicity in the less preferred (off diagonal) combinations as well. The F/C papers often present the off-diagonals (e.g., Davis, MacNeilage, & Matyear, 2002), but the only analysis presented is that there are fewer ratios larger than 1 in the off-diagonals; no analysis of the patterning in the off-diagonals is attempted. However, our modeling results show that there is systematicity in the off-diagonals that is well-predicted by Articulatory Phonology (AP), which can calculate the synergy between any two phones. Our modeling of F/C did not support the assumptions of that model

(Nam, Giulivi, Goldstein, Levitt, & Whalen, submitted). Because of this systematicity, we will expand the investigation to the off-diagonals here.

In order to test whether spoken adult language could be the source for the CV preferences in babbling, we examined CV combinations in spoken corpora for the three languages used in our own study of babbling: English, French and Mandarin (Giulivi, 2007; Giulivi, et al., 2011). The adult spoken corpora we examined are larger than those used in Janson's study, which should provide a more stable indication of the existence of CV combination preferences in token counts of adult language. In addition to our spoken corpus count, we also compare dictionary counts for English and Mandarin, to allow us to study the patterning of all nine CV combinations rather than just the three preferred patterns.

2 Testing favored CV co-occurrences in two spoken English corpora

The primary measure that has been used to demonstrate the CV combination preferences is a ratio of the observed CV combinations to the expected number based on the overall occurrence of C and V independently (e.g., Davis & MacNeilage, 1995). In previous work, all vowels were classified as Front, Central or Back, and stop and nasal consonants were classified as Labial, Coronal or Velar. Fricatives were excluded from the babbling studies due to their low rate of occurrence, as were the liquids [r, l]. The glides [w, j] were included in their counts, with [w] counting as a labial. To maintain comparability, the same criteria were used with the adult corpora.

2.1 Method

The consonants that were included were the stops, nasals and glides (i.e., excluding affricates, fricatives and liquids) and were grouped, according to place of articulation, into labials [p, b, m, w], coronals [t, d, n, j] and velars [k, g, ŋ]. Vowels

were grouped with reference to the front-back dimension of the vocal tract, into central [ʌ, ə, a], front [i, ɪ, e, ε, æ] and back [u, ʊ, o, ɔ].

The first English corpus was the Switchboard database of English spoken by 542 different talkers (Godfrey & Holliman, 1997). Each speaker produced short monologues in response to one of 70 prompts on various topics (e.g., taxes, pets, movies). This resulted in 3,068,137 words available in the transcriptions, with 1,474,728 CV syllables that met the criteria for inclusion here.

The second English corpus was the Buckeye database (Pitt, et al., 2007), containing recordings and transcriptions of interviews with 40 American English speakers of different ages from a single dialect area (Central Ohio). A total of about 300,000 words were transcribed, resulting in 143,756 CV syllables that met our criteria.

2.2 Results

Table 1 presents the ratios for the Switchboard corpus, and Table 2, for the Buckeye corpus. In both cases, only one of the three diagonals is greater than 1. This finding is quite different from the previously reported results for both babbling and English type (dictionary) counts.

Tables 1 and 2 about here.

The diagonals have been shown to represent only a portion of the systematicity present in the data (Nam, et al., submitted). A fuller account can be achieved with a Pearson correlation between all nine cells of one measure with another. With such an analysis, the results from the two spoken English corpora are quite similar; a correlation of the ratios gives an r of .96 ($p < .001$). The correlations

with the original ratios from the babbling data (Davis & MacNeilage, 1995), on the other hand, are not only small in magnitude but negative as well ($r = -0.23$ for the Switchboard, -0.11 for the Buckeye, neither significant). The correlations with the English babblers of our own study (Giulivi, et al., 2011) are also negative and large enough to be significant ($r = -0.77$ for the Switchboard, -0.70 for the Buckeye, $p < .05$ for both).

Because the previously reported dictionary counts were presented only graphically, we recalculated dictionary counts from two sources so that we could explore all nine cells of the CV combinations with those of the dictionary data. The first was the Shorter Oxford English Dictionary (SOED) as made available in the MRC Psycholinguistics Database (Coltheart, 1981), the same one used by MacNeilage et al. (2000). It contains phonetic transcriptions for 38,420 words (58,862 CV syllables used for our analysis). The second dictionary was the Carnegie Mellon University (CMU) Pronouncing Dictionary (CMU_Speech_Group, 2007), which contains over 112,131 words for our search, yielding 133,797 CV syllables for analysis. Tables 3 and 4 present the ratios of observed to expected occurrences, presented as before.

Insert Tables 3 and 4 about here.

The ratios are similar to those reported by MacNeilage et al. (2000). There are discrepancies in the absolute magnitudes for the SOED, even though that was their source as well as ours. We do not have an explanation for this difference. The correlation of the nine ratios from the two dictionaries was good ($r = 0.89$, $p < .01$), showing that the difference in size and dialect had little effect on the ratios.

Although the diagonals show general agreement between the babbling (token) and dictionary (type) counts, the correlation of all nine cells does not. Neither dictionary was significantly correlated with either the MacNeilage and Davis data ($r = 0.02$ for SOED and 0.24 for CMU) or with our English results ($r = -0.09$ for SOED and 0.05 for CMU).

The correlations between the nine ratios from the spoken corpora to the corresponding dictionary-based ones were not significant. Neither spoken corpus was significantly correlated with the SOED ($r = 0.40$ for both), with similar results for the CMU ($r = 0.37$ for the Switchboard corpus, $r = 0.31$ for the Buckeye, n.s.). These low correlations indicate that the frequencies of certain words affect the rate at which the combinations preferred in babbling appear in adult speech. Notably, such words as “because” and “got” (velar-central), “you” and “know” (coronal-back), and “we” and “well” (labial-front) are quite common and occur in non-preferred cells.

2.3 Discussion

The pattern of preferred CV combinations in English babbling is significantly different from the CV combination preferences in the ambient spoken language. This finding suggests that universal tendencies have a stronger influence on babbling than does the language environment.

In addition, the ratios obtained from spoken English (token counts) and those obtained from dictionaries (type counts) were not significantly correlated, and it is only the dictionary counts that provide evidence in support of the preferred combinations in babbling. Thus, even for the relatively weak comparison of the diagonal ratios, only one of the three CV combinations preferred in babbling was preferred in the adult token counts (the labial-central combination). This result stands in contrast to the diagonals of the type counts, which were generally supportive of the

babbling pattern, as previously reported (2 of 3 diagonals for the SOED, all three for the CMU dictionary).

On our more stringent test of comparing all nine cells of the spoken corpora with those of the dictionary and babbling results (thus including secondary regularities), the correlations were either nonsignificant or, in two cases, negatively correlated. Thus it appears that the frequency of usage in spoken English fails to reflect the slight tendency for word types to exhibit the preferred combinations.

3. CV favored co-occurrences in two French corpora

Another of the languages in our babbling study was French (Giulivi, et al., 2011), so a further comparison was made with two corpora of spoken French.

3.1 Method

The first corpus was that of Tubach and Boë (1990). They aggregated three corpora: one of presentations and discussions at a conference, one of conversations between educated adults, and one of sixteen brief conversations among adults, which yielded 3,265 CVs for analysis. The second corpus was the Lexique 3 database, version 3.2 (c.f. New, Pallier, Brysbaert, & Ferrand, 2004). Although a large part of this dataset is based on written texts, we analyzed the segments that were taken from subtitles for films. These were thus scripted dialogs, but were, in general, attempts at replicating spoken language. That this is a reasonable approximation to spoken language is supported by a recent study for Dutch, in which it was found that frequencies based on subtitles were, in fact, capable of explaining more variance of lexical decision reaction times than counts based on published texts (Keuleers & Brysbaert, 2010). Here, the same criteria for classifying the syllables were used as before, and ratios were calculated in the same way. 184,012 syllables were available for analysis.

3.2 Results

Table 5 shows the results for the Tubach and Boë data. Here, the diagonals are in good agreement with the preferences in babbling; indeed, they are noticeably larger than those found in babbling.

Table 5 about here.

Table 6 shows the results for the French subtitles. Similarly to Table 5, the diagonals are noticeably larger than those found in babbling.

Table 6 about here.

The correlations for the full set of nine cells were calculated for the two French corpora and the babbling data. The corpora correlated significantly with each other ($r = 0.83$, $p < .01$), as did each with the MacNeilage and Davis babbling data ($r = 0.72$ and 0.76 , $p < .05$, for the Tubach/Boë and Lexique, respectively). The Tubach/Boë also correlated with our babblers in the French environment ($r = 0.78$, $p < .05$), but the Lexique did not ($r = 0.58$, n.s.). Neither corpus correlated significantly with our babbling data from the other two environments (English and Mandarin).

In order to have a full set of ratios for a type countⁱ, we analyzed the French subtitles by type rather than token (Table 7). This is similar to a dictionary count, but applies to the same material as that of the token count. Again, the values on the diagonals are greater than 1.0, demonstrating the same pattern of preferences observed in the babbling data, but of a greater numerical magnitude. The Tubach/Boë ratios were marginally correlated with this type count ($r = 0.66$, $p = 0.055$), and the Lexique token ratios were significantly correlated with the type counts from that same

data ($r = 0.87$, $p < .01$). This last correlation is no doubt higher than would be expected to obtain with a type count based on an unrelated set of texts.

Table 7 about here.

3.3 Discussion

The ratios for the French film subtitles are somewhat higher along the diagonal than those in the babbling data (Giulivi, et al., 2011) and in the French dictionary data (1.2, 1.6 and 1.6 for the diagonals; MacNeilage et al., 2000). They are thus in general accord with the previously reported results. The correlations with the babbling results (MacNeilage and Davis's and our French-learning infants) are positive. In isolation, we would have taken this finding as evidence of early attunement to the language environment in babbling, especially because the correlations between the spoken corpora and our English-learning infants were not significant. Set against the English results, however, it appears that this correlation is due to a greater agreement between spoken French and the universal pattern than any direct attunement. Adding a third case, here, Mandarin, will help elucidate this issue.

4. CV favored co-occurrences in a Mandarin corpus

The final language in our babbling study was Mandarin (Giulivi, et al., 2011), so a comparison was made with a corpus of spoken Mandarin. Because previous dictionary counts did not include Mandarin, we provide a count based on a searchable internet dictionary of Mandarin (CC-CEDICT, 1997).

4.1 Method

The Chinese Annotated Spontaneous Speech (CASS) version 3.2 (Li, et al., 2000) database was used here. It was designed to explore the phonetic realization of

Mandarin words in spontaneous speech and contains narrow transcriptions of 3 hours of speech from 7 talkers and includes 50,782 syllables. As before, those syllables with fricatives, affricates and liquids for onsets were excluded, leaving 24,766 syllables. The consonants were classified as labial (transcribed as b, p, m, w), coronal (t d n y) or velar (k g). The vowels were classified as front (transcribed as i, ei, ü, ia, ie, iao, iu), central (a, ao, e) or back (o, ou, u, ua, uo, uai, ui). Note that the "i" or "u" onglide in some of the vowel combinations is treated as a vowel by the phonotactics of Mandarin, not as a consonant; we thus classified those as vowels.

Our comparison for the dictionary count was taken from the on-line Chinese-English Dictionary - (CEDICT). For this, all 84,303 syllables with pinyin transcriptions were examined. Those syllables with fricatives, affricates and liquids for onsets were excluded, leaving 32,715 syllables. The same criteria as CASS corpus were used for CV classification.

4.2 Results

Table 8 shows the results for the spoken corpus of Mandarin. The diagonals resulted in two that were greater than 1 (the coronal-front and velar-back combinations). The labial-central ratio was well below 1. Thus there is limited evidence of the preferred syllables in spoken Mandarin.

Table 8 about here.

The correlations between the Mandarin token ratios and the babbling ratios were not significant ($r = 0.04$ for the MacNeilage/Davis data, -0.08 for our Mandarin-learning infants). Neither of the correlations with our other two language groups was significant.

Table 9 shows the results for the dictionary of Mandarin. Here again, two of the three values on the diagonals were greater than 1, indicating some evidence for the preferred combinations in lexical counts. In this case, the labial-central value was relatively close to 1.

Table 9 about here.

The pattern of ratios for the spoken (token) corpus correlated with the dictionary (type) ratios ($r = 0.73, p < .05$). As could be expected, then, the dictionary ratios did not correlate with the babbling results ($r = 0.03$, n.s., for MacNeilage/Davis, -0.36 , n.s., for our Mandarin-learning infants).

4.3 Discussion

Two of three diagonal ratios were greater than 1 in both Mandarin spoken corpora, whereas the other was less than 1. Thus there is limited support for the persistence of CV preferences into adult spoken language for Mandarin. However, the ratios for Mandarin were much more similar between the dictionary and spoken corpora than were English and French. This pattern stands in contrast to the marginal results for French and the lack of a correlation for English.

Neither the type nor token ratios from adult Mandarin correlated with the babbling results.

5. General discussion

Previous work had found that the CV combinations that are preferred in babbling are similar to those combinations as they occur in type (dictionary) counts for a variety of languages. In the present study, we instead focused on spoken corpora

as being more representative of the language that infants would hear. The patterns of preferred CV combinations were similar for babbling and spoken corpora only for French. For Mandarin, there was no correlation, and there was a negative correlation for English. This indicates that the CV preferences are likely to have a biomechanical source since they appear in the babbling in all these language environments, that is, regardless of the frequency of the CVs in the input to the infant. In addition, we found that the previously reported presence of similar preferences for these three combinations in dictionary (type) counts for various languages (MacNeilage, et al., 2000) was supported to the same extent (10 out of 12 comparisons) in the type counts examined. However, the connection to the babbling data was less consistent. The correlation between all nine CV combinations of the babbling (token) and dictionary (type) counts was not significant for any of the three languages (English, French and Mandarin), even though simulations of the AP biomechanical account had been found to correlate with all nine cells (Nam, et al., submitted). Perhaps the strongest indication of such further systematicity in the off-diagonals is the consistently low ratios for the velar-front combination, a particularly challenging articulatory combination.

The likelihood that there is a biomechanical source for the CV preferences in babbling leads to the expectation that the pattern will be universal. Although the CV preferences have been found in most of the languages that have been investigated for this pattern (English, French, Mandarin (though cf. Chen & Kent, 2005), Italian, (Zmarich & Miotti, 2003), Korean (S. Lee, Davis, & MacNeilage, 2007)), there have been reports that Dutch babbling does not follow the pattern (Schauwers, Gillis, & Govaerts, 2008; Zink, 2005). These studies found that children in Dutch-speaking environments had the expected preference for coronal/front combinations, but that

labials occurred more often with back vowels. Velar consonants did not have co-occur with any vowel class at greater than chance levels. Part of the discrepancy is that the vowel in utterances such as "papa" was counted as back (Zink, 2005), even though particular instances might have been central. Another part of the discrepancy is that individual ratios were tested for significance. The labial/central combination for the normal hearing group of Schauwers et al. (2008), for example, gives a ratio of 1.05; this is above 1, as expected, but it is not significantly so by their measure. The velar/back combination, on the other hand, has a (nonsignificant) ratio of 0.8, making Dutch at best one of the languages that has two out of three preferred combinations. In short, there is too little data to be confident that the pattern is universal, but the plausibility of the AP biomechanical account and the appearance of the pattern despite contrary frequencies in the ambient language input (English and Korean) leads to the expectation that most languages will conform to the pattern.

The appearance of the preferred CVs in the adult dictionaries remains unexplained. The F/C account does not predict the adult pattern, and it indeed posits a second mechanism. MacNeilage et al. (2000) assert that the maintenance of the CV preferences in adult language "probably means that the patterns have been present since the origin of speech in hominids" (p. 158). Similarly, Davis et al. (2002) state, "The common occurrence of these results in infants, languages and the proto-word corpus suggests a fundamental status for these CV co-occurrence constraints in languages" (p. 94). Both of these explanations rely on continuity between babbling and language, whereas the emergence into adult speech is claimed to occur when the infant "escapes frame dominance" (Davis & MacNeilage, 1995:1208). The persistence of the CV preferences in adult language is, then, unexpected. Even the more successful biomechanical account, AP, does not necessarily predict that CV

synergy will be represented in the dictionary. Overall usage of words is clearly not dictated by these synergies, as seen especially in the English spoken corpora. Adults have greater control over their articulators than infants do, and thus they are not as subject to the preferences that arise biomechanically from the combinations of certain gestures. We might expect, however, that there would be circumstances in which such small preferences could have an effect, such as in speeded production of nonsense words or judgments of which string of Cs and Vs would make a better word. Such small effect may be more powerful in rare words, leading to an overall preference for the synergistic CV combinations in the dictionary as a whole. Still, if a language's lexicon did not exhibit these preferences, we would not take that as evidence that the synergies were different in that language, only that they did not happen to influence the lexicon.

The relationship between the type and token counts for the adult language is weakly positive: The Mandarin corpus was positively correlated, while the French and English corpora were uncorrelated. The English results, in particular, seem to be due to a sizable number of very frequent words that contain non-favored combinations. The Mandarin correlation may also be due to a design decision that was made to maintain comparability with the babbling results. For babbling, fricatives and affricates were excluded from the C counts due to their low frequency. Given that Mandarin has a phonological process of affrication of alveolars before high front vowels (Chao, 1968:21), this resulted in a great many exclusions from what would have been the preferred coronal-front combination. It may be that including those would change the pattern of type and token frequencies so that they, like the French and English, would be uncorrelated.

In summary, we have found evidence that the preferences found for certain CV combinations emerge in babbling even if the adult language does not show the same pattern. This indicates that the previously proposed biomechanical explanations are necessary, because the infants are not always imitating the input. The persistence of these CV preferences in adult lexicons remains to be explained, given that adult spoken frequencies often depart from those patterns.

References

- Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, 72, 416-430.
- Boysson-Bardies, B. d., Hallé, P. A., Sagart, L., & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16, 1-17.
- CC-CEDICT (1997). CC-CEDICT. <http://cc-cedict.org/wiki/>
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Chen, L.-M., & Kent, R. D. (2005). Consonant-vowel co-occurrence patterns in Mandarin-learning infants. *Journal of Child Language*, 32, 507-534.
- CMU_Speech_Group (2007). CMU pronouncing dictionary. <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.
- Davis, B. L., & MacNeilage, P. F. (1994). Organization of babbling: A case study. *Language and Speech*, 37, 341-355.

- Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, 38, 1199-1211.
- Davis, B. L., MacNeilage, P. F., & Matyear, C. L. (2002). Acquisition of serial complexity in speech production. A comparison of phonetic and phonological approaches to first word production. *Phonetica*, 59, 75-107.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181-195.
- Giulivi, S. (2007). *Vowels and consonants favored co-occurrences in language development*. Unpublished Ph.D. dissertation, University of Florence.
- Giulivi, S., Whalen, D. H., Goldstein, L. M., Nam, H., & Levitt, A. G. (2011). An Articulatory Phonology account of preferred consonant-vowel combinations. *Language Learning and Development*, 7, 202-225.
- Godfrey, J. J., & Holliman, E. (1997). Switchboard-1 Release 2. from Linguistic Data Consortium:
- Honda, K. (1996). Organization of tongue articulation for vowels. *Journal of Phonetics*, 24, 39-52.
- Jakobson, R. (1968). *Child language, aphasia and phonological universals*. The Hague: Mouton.
- Janson, T. (1986). Cross-Linguistic Trends in the Frequency of CV Sequences. *Phonology Yearbook*, 3, 179-195.
- Keuleers, E., & Brysbaert, M. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42, 643-650.

- Lee, S., Davis, B. L., & MacNeilage, P. (2007). "Frame Dominance" and the serial organization of babbling, and first words in Korean-learning infants. *Phonetica*, 64, 217-236.
- Lee, S. A. S., Davis, B. L., & MacNeilage, P. F. (2010). Universal production patterns and ambient language influences in babbling: A cross-linguistic study of Korean- and English-learning infants*. *Journal of Child Language*, 37, 293-318.
- Li, A., Zheng, F., Byrne, W., Fung, P., Kamm, T., Liu, Y. S., Zhanjiang, Ruhi, U., Venkataramani, V., & Chen, X. X. (2000). *CASS: A phonetically transcribed corpus of Mandarin spontaneous speech*. Paper presented at the ICSLP-2000, Beijing.
- Lindblom, B. E. (1986). Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental phonology* (pp. 13-44). Orlando, FL: Academic.
- Lindblom, B. E. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403-439). Dordrecht: Kluwer Academic Publishers.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21, 499-546.
- MacNeilage, P. F., & Davis, B. L. (1990). Acquisition of speech production: Frames, then content. In M. Jeannerod (Ed.), *Attention and performance XIII* (pp. 453-476). Hillsdale, NJ: Lawrence Erlbaum Associates.
- MacNeilage, P. F., & Davis, B. L. (2011). In defense of the "Frames, then Content" (FC) perspective on speech acquisition: A response to two critiques. *Language Learning and Development*, 7, 234-242.

- MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (2000). The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development*, 71, 153-163.
- Maddieson, I. (1984). *Patterns of sounds*. New York: Cambridge University Press.
- Maddieson, I. (1997). Phonetic universals. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 619-639). Oxford: Blackwell.
- Maddieson, I., & Precoda, K. (1992). Syllable structure and phonetic models. *Phonology*, 9, 45-60.
- Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech*, 21, 221-241.
- Nam, H., Giulivi, S., Goldstein, L. M., Levitt, A. G., & Whalen, D. H. (submitted). Computational simulation of CV combination preferences in babbling.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods Instruments & Computers*, 36, 516-524.
- Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Oshima-Takane, Y. (1988). Children learn from speech not addressed to them: the case of personal pronouns. *Journal of Child Language*, 15, 95-108.
- Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). Buckeye corpus of conversational speech (2nd release). from Columbus, OH: Department of Psychology, Ohio State University (Distributor): [www.buckeyecorpus.osu.edu]

- Rousset, I. (2004). *Structures syllabiques et lexicales des langues du monde: Données, typologies, tendances universelles et contraintes substantielles*. Université Grenoble, Grenoble.
- Schauwers, K., Gillis, S., & Govaerts, P. J. (2008). The characteristics of prelexical babbling after cochlear implantation between 5 and 20 months of age. *Ear and Hearing*, 29, 627-637.
- Tubach, J. P., & Boë, L.-J. (1990). *Un corpus de transcription phonétique (300.000 phones): Constitution et exploitation statistique*. Paris: Telecom Paris.
- van de Weijer, J. (2001). *How much does an infant hear in a day?* Paper presented at the GALA (Generative Approaches to Language Acquisition).
- Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From babbling to speech: A re-assessment of the continuity issue. *Language*, 61, 397-445.
- Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, 17, 323-333.
- Whalen, D. H., Giulivi, S., Goldstein, L. M., Nam, H., & Levitt, A. G. (2011). Response to MacNeilage and Davis and to Oller. *Language Learning and Development*, 7, 243-249.
- Whalen, D. H., Levitt, A. G., & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language*, 18, 501-516.
- Zink, I. (2005). De verwerving van de klankproductie tijdens de brabbelperiode bij vier Vlaamse kinderen. *Logopedie*, 18, 13–20.

Zmarich, C., & Miotti, R. (2003). The frequency of consonants and vowels and their co-occurrences in the babbling and early speech Italian children *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1947-1950). Barcelona.

English—Switchboard			
C \ V	Front	Central	Back
Coronal	0.83	0.87	1.51
Labial	1.25	1.12	0.41
Velar	0.90	1.11	0.94

Table 1. Ratios of observed to expected frequencies generated for spoken English (Switchboard database).

English—Buckeye			
C \ V	Front	Central	Back
Coronal	0.79	0.67	1.67
Labial	1.26	1.31	0.29
Velar	0.95	1.27	0.77

Table 2. Ratios of observed to expected frequencies generated for spoken English (Buckeye database).

English--SOED			
C \ V	Front	Central	Back
Coronal	1.07	0.90	1.04
Labial	1.04	0.99	0.89
Velar	0.74	1.28	1.08

Table 3. Ratios of observed to expected frequencies generated for a dictionary of English (SOED database).

English--CMU			
C \ V	Front	Central	Back
Coronal	1.08	0.92	1.02
Labial	1.01	1.03	0.88
Velar	0.80	1.12	1.17

Table 4. Ratios of observed to expected frequencies generated for a dictionary of English (CMU database).

French—Spoken, Token counts			
C \ V	Front	Central	Back
Coronal	1.21	0.70	0.79
Labial	0.74	1.61	0.85
Velar	0.96	0.57	1.86

Table 5. Ratios of observed to expected frequencies generated for a corpus of French lectures and conversations, token count.

French—Subtitles, Token counts			
C \ V	Front	Central	Back
Coronal	1.27	0.42	0.91
Labial	0.61	1.59	1.17
Velar	0.57	1.00	1.75

Table 6. Ratios of observed to expected frequencies generated for a corpus of French subtitles in films (Lexique 3 database), token count.

French—Subtitles, Type counts			
C \ V	Front	Central	Back
Coronal	1.20	0.74	0.77
Labial	0.83	1.10	1.13
Velar	0.44	1.10	1.68

Table 7. Ratios of observed to expected frequencies generated for a corpus of French subtitles in films (Lexique 3 database), type count.

Mandarin--Spoken			
C \ V	Front	Central	Back
Coronal	1.30	0.96	0.79
Labial	1.02	0.86	1.28
Velar	0.09	1.35	1.15

Table 8. Ratios of observed to expected frequencies generated for a corpus of spoken Mandarin (CASS database).

Mandarin--Dictionary			
C \ V	Front	Central	Back
Coronal	1.17	1.03	0.81
Labial	1.25	0.94	0.83
Velar	0.01	1.05	1.87

Table 9. Ratios of observed to expected frequencies generated for a dictionary of Mandarin (CEDICT database).

ⁱ We did not have access to a full dictionary for French.