J. Child Lang. **32** (2005), 507–534. © 2005 Cambridge University Press
doi:10.1017/S0305000905006896 Printed in the United Kingdom

Consonant–vowel co-occurrence patterns in Mandarin-learning infants*

LI-MEI CHEN
National Cheng Kung University, Taiwan, R.O.C.

AND

RAYMOND D. KENT
University of Wisconsin-Madison

(*Received* 21 *May* 2003. *Revised* 16 *July* 2004)

ABSTRACT

Most studies on CV co-occurrence patterns in early phonetic development have been based on Indo-European languages. Data from infants learning Mandarin, which has a substantially different phonological system from Indo-European languages, can confirm or refute the findings of previous studies, thus shedding further light on the theoretical bases of CV association. Spontaneous vocalizations were recorded in 45-minute sessions for each of 24 Mandarin-learning infants aged 0;7 to 1;6. In addition, the speech production of 24 caregivers was audio-recorded during their natural daily interactions with the infants at home. Both infants' vocalizations and adults' speech were transcribed and analysed for consonant–vowel co-occurrence patterns. These patterns were similar to those in other language groups, but language-specific patterns were evident by 1;0. Combinations of alveolars + front vowels and velars + back vowels confirm Davis & MacNeilage's (1990, 1995) frame-then-content theory and Clements's (1991) model of unified features for consonants and vowels. However, our finding of a language-specific pattern (labials + back vowels) suggests the need to reexamine the 'pure frame' of Davis & MacNeilage and Clements's grouping of labials and rounded vowels.

[*] We are grateful to the infants and their families who participated in both this study and the pilot study. Without their involvement, this work would not have been possible. Their hospitality made the whole recording procedure an enjoyable experience. In revising the manuscript, we appreciate the insightful and detailed comments from the reviewers. Address for correspondence: Raymond D. Kent, Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, Wisconsin, 53705, USA. e-mail: Kent@waisman.wisc.edu

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

INTRODUCTION

CV association in early phonetic development has been well studied, with various theoretical approaches standing in interesting contrast. For example, theories on crosslinguistic phonotactic constraints (with either acoustic or articulatory bases) explore the bases and patterns of CV affiliation in speech development. In addition, frame dominance, task dynamic, hyperspeech and hypospeech, and feature geometry models have been advanced to account for the patterns of phoneme sequences in developing systems. Some of these models are expressed in phonetic terms, others in phonological terms, and some in terms that do not fit in either traditional phonetic or phonological frameworks. These important differences carry implications for the role of babbling in acquiring early words and developing a phonological system.

However, this territory is also treacherous because one should not assume that identifying a given sound in infant vocalizations provides evidence of that sound's status in a phonemic inventory or a phonological system of any kind. Some argue that phonology *per se* is not established until a child has an expressive vocabulary of at least 50 words, for only at that point is the child motivated to seek phonological principles. The following review considers first the theoretical proposals regarding CV affiliation and second, the empirical evidence obtained in studies of infant sound production.

Theoretical bases

Kawasaki-Fukumori (1992) claimed that commonly occurring syllabic structures, especially the CV affiliation pattern, have perceptual/acoustical bases. Considering a large number of the languages of the world, Kawasaki-Fukumori summarized three disfavoured phonotactic patterns, which he suggested are generalizable to all languages: (1) labial consonants + /w/, round vowel, or back vowel; (2) alveolar consonants + /j/, or front vowel; and (3) velar consonants + round vowel, front vowel, or high vowel. In addition, there were two favoured patterns: (1) /w/ or labialized consonants + front or low vowels; and (2) /j/, dental, alveolar, or palatalized consonants + back or low vowels. These phonotactic constraints reflect the hypothesis that acoustical patterns with high perceptual salience (large spectral change within a syllable) are generally more efficient in perception, and thus favoured, while sequences with small spectral changes are disfavoured. However, empirical data used to support Kawasaki-Fukumori's arguments on the world's languages provide sufficient evidence for only two generalizations: (1) /w/ or labialized consonants + rounded vowels is disfavoured; and (2) /j/ or palatalized consonants + front vowels is disfavoured. The other general statements on disfavoured patterns of CV

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

co-occurrence are based on very limited data and several other studies on CV association provide evidence contrary to Kawasaki-Fukumori's general summaries.

In an investigation of five languages (Finnish, Turkish, Latin, Lativian, and Setswana), Janson (1986) found the following CV association patterns: (1) dentals and alveolars /n, s, d, l/ tend to be followed by front vowels; (2) there are no clear general trends for labials, except for the association of /b, p/ and back vowels; and (3) velar consonants occur least frequently with /k/, with a back vowel being the only clear association trend. These association trends (based on articulatory constraints on crosslinguistic CV sequences) contradict Kawasaki-Fukumori's acoustic-based generalizations. Janson (1986: 193) indicated that these crosslinguistic trends in CV combinations involve articulators that 'do not have to make extensive movements from the consonant gesture to the vowel gesture.' This concept is usually referred to as ease of articulation. The articulatory-based account of CV sequences suggests that the reason for the associations of dental + front vowel and velar + back vowel is that consonants and vowels share similar places of articulation in each set. However, the only conclusion in Janson's study with sufficient supporting data is that concerning the association pattern of alveolars with front vowels. The other generalizations were based on only five European languages and do not provide sufficient evidence to affirm universal trends of CV affiliations.

Taking the theoretical framework of 'ease of articulation' one step further, Davis & MacNeilage (1990, 1995) and MacNeilage (1996) proposed another articulatory-based model of speech production, frame-then-content theory, and suggested a neurobiological basis for CV association patterns found in developmental data. They asserted that both babbling and early speech can be seen as rhythmic alternation between a closed (consonant-like sound) and an open (vowel-like sound) mouth accompanied by phonation. This fundamental motoric pattern in babbling and early speech can be accounted for mainly by mandibular oscillation with passive movement of the tongue and soft palate (in nasal environment).

This idea is similar to Hodge (1989) on infant babbling that jaw movement plays a major role in early syllables, with the tongue riding passively on the jaw movement. MacNeilage (1996, 1997) stated that in symmetrical nasal environments there are at least four basic mandibular oscillation patterns (i.e. frames for producing early CV syllables): (1) pure frame–labials with central vowels, (2) front frame–alveolars with front vowels, (3) back frame–dorsals with back vowels, and (4) nasal frame–nasal resonance of vowels (in formant frequencies of low vowels and front vowels) (Matyear, MacNeilage & Davis, 1998). These frames were thought to offer initial sets of contexts at early stages of development and it is assumed that context-independent segmental components (content) emerge later.

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

Addressing nature and nurture factors in development, MacNeilage (1996) suggested that both motor control in producing frames and the input of syllable structures from the ambient language contribute to development, in that one provides a context for the other. That is, according to MacNeilage (1997), in frame-then-content theory, the development of infant vocalizations is the combined result of the general linguistic input, self-input, and the developing output propensities. However, as de Boysson-Bardies (1993) pointed out, this theory, in fact, emphasizes articulatory motor control, and the role of linguistic input in shaping early phonological forms has not been adequately incorporated.

Instead of segments, phonemes or features, Browman & Goldstein (1992) proposed articulatory gestures as the basic phonological units. These gestures in particular were suggested by Browman & Goldstein for describing articulatory routines or recurring patterns of well-practiced units in babbling and early speech. According to Browman & Goldstein, gestures are advantageous for describing the prelinguistic system for two reasons. First, gestural description (e.g. /b, d, g/ as gross constriction gestures of the lips, tongue tip, and tongue body, respectively, /h/ as a laryngeal widening gesture) was found to be more adequate and reliable than phonetic transcription, which often inadequately imposes units in adult phonological systems onto prelinguistic vocalizations. Second, because gestural units with intrinsic duration may overlap in time, many cases of co-occurrence patterns or assimilation can be derived directly from gestural overlap, and there is no need to posit phonological adjacency constraints (e.g. feature spreading). Browman & Goldstein suggested that children's early productions are a small number of 'dynamically stable patterns' consisting of inadequate and imprecise gestural overlapping as undifferentiated units. Children gradually learn to differentiate whole units into individual components and to coordinate these primitive gestures in producing early syllables. If Browman & Goldstein's ideas of differentiation and coordination are extended to account for constraints on phoneme sequences, CV co-occurrence patterns found in early infant vocalizations may result directly from natural gestural overlapping before independent controls for individual gestures are developed. In other words, this model predicts that consonants and vowels that share the same places of constriction tend to co-occur in early CV syllables. Later in development, children acquire language-specific CV association patterns out of the original patterns of same-place articulation.

Lindblom's (1990) Hyper and Hypo (H & H) Theory suggested that possible speech sounds and sound combinations are derived from the process of speaker-listener interaction, and are constrained by the principles of sufficient discriminability (hyperspeech) and articulatory simplicity (hypospeech). A physiologically based numerical model of computational simulation for CV-like syllables (Lindblom, MacNeilage, &

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

Studdert-Kennedy, 1983), based on principles of both low articulatory cost and high discriminability, was used to generate 15 CV syllables:

bi	bɛ	ba	bo	bu
di	dɛ	da	do	du
g+i	g+ɛ	ga	go	gɨ

(/g+/: palatal variation of /g/. These phonetic symbols are based on Lindblom *et al*., 1983)

In these CV syllables, three voiced stops /b, d, g/ and five vowels /i, ɛ, a, o, u/ are repeatedly used. This model predicts the following CV co-occurrence patterns in the world's languages: (1) labials accompany either front or back vowels; (2) dentals accompany either front or back vowels; (3) palatals accompany front vowels; and (4) velars accompany nonhigh, nonfront vowels. Lindblom's model did not suggest any explicit assimilation rules for sound sequences of CV syllables in a developmental system. Instead, Lindblom proposed that early CV syllable patterns represent an unplanned emergence of adaptation to both articulatory and perceptual constraints during the acquisition of increasing amounts of lexical items. Lindblom's emergent and self-organizing model plausibly explains how children rapidly acquire a phonological system and a large lexicon from the ambient language while facing an 'invariance problem' from input speech. However, it is not clear how this model accounts for early CV sequences in pre-lexical babbling. This theory requires more exploration to account for the very early stage of the phonological system preceding extensive lexical development.

Apart from models with articulatory, perceptual and self-organizing bases for CV association patterns, various models of feature geometry from non-linear phonology provide distinct perspectives regarding CV association based on phonological patterning of the world's languages. These models have been developed and revised based on evidence from assimilation and dissimilation rules and obligatory contour principles in the world's languages. Data from phonological acquisition are rarely used to test or establish the numerous proposals of feature geometry. However, since various models of feature geometry differ both in the way place features of consonants and vowels are organized and in the planar relationship among these features, as suggested by Gierut, Cho & Dinnsen (1993), CV interaction from developmental data is one form of crucial evidence (in addition to evidence from fully-developed languages) for evaluating the different perspectives of feature geometry.

In Halle-Sagey's (Halle, 1992) articulatory-based model, place features of consonants and vowels are organized under three articulator nodes (Figure 1).

One major characteristic of this model is that the three vowel features, [high], [low], and [back], are all linked under the dorsal node. This

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

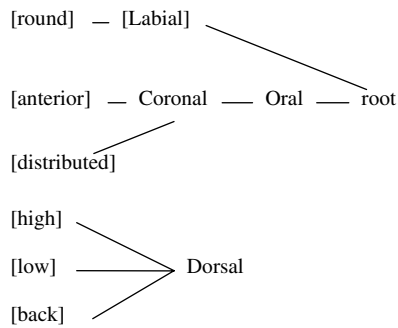


Fig. 1. Halle-Sagey's feature geometry (from Halle, 1992, only the portion of the geometry relevant to the current discussion is shown in this figure).

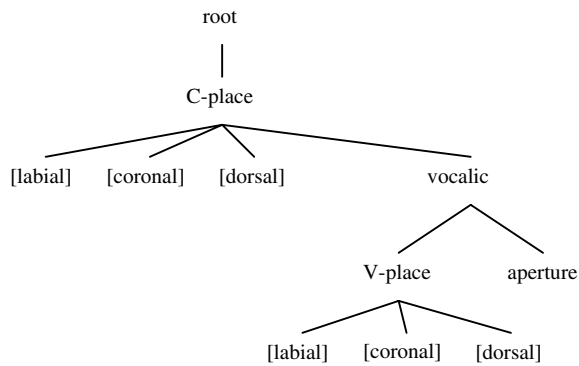


Fig. 2. Clements's (1991) feature geometry (only the portion of the geometry relevant to the current discussion is shown in this figure).

characteristic leads to the prediction that dorsal consonants and all vowels will undergo a phonological process as a natural class. Furthermore, if CV interaction in developmental data can be seen as an assimilation rule that spreads place features of vowels to consonants, then this model predicts that all vowels in early CV syllables will combine with dorsal consonants. However, this prediction has not been found in available developmental data.

In Clements's (1991) constriction-based model, four articulatory features – labial, coronal, dorsal, and radical – are used to specify place of articulation in both consonants and vowels (Figure 2).

Through assimilation rules that apply among adjacent CV segments, this model makes three predictions about early CV syllables: (1) labial consonants tend to occur with rounded vowels; (2) coronal consonants tend to accompany front vowels; and (3) dorsal consonants (e.g. velar and labiovelar

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

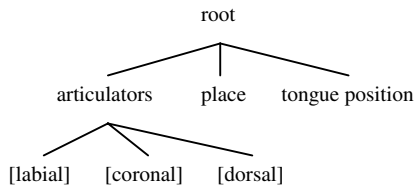


Fig. 3. Lahiri & Evers's (1991) feature geometry (only the portion of the geometry relevant to the current discussion is shown in this figure).

consonants) tend to co-occur with back vowels. These predictions were also found in numerous observations of CV co-occurrence patterns in early phonological systems.

Lahiri & Evers's (1991) model of feature geometry proposes complete overlapping of C-place and V-place planes. In this model, place features of both consonants and vowels are assigned to the same articulatory node, which is immediately dominated by the place node as shown in Figure 3.

In Lahiri & Evers's model, all the places of articulation of consonants and vowels can be characterized by the same set of articulatory features (labial, coronal and dorsal). According to the Locality Condition, any two adjacent segments, including adjacent consonants and vowels, are free to interact. Therefore, this model generates predictions similar to those of Clements's, except that Lahiri & Evers's model predicts an extra association pattern for velar consonants (i.e. velars with round vowels). Comparison of these three models reveals that the feature geometries of Clements's (1991) and Lahiri & Evers's (1991) models better account for CV associations found in children's early CV syllables and yield the same predictions for the association patterns.

The application of feature geometry to early CV association patterns raises several critical issues and offers several hypotheses for studying phonological acquisition. Because vowels are acquired and stabilized earlier than consonants, it is assumed that early CV association occurs because the place features of consonants are generally underspecified, and consonants acquire place features from adjacent vowels through assimilation rules. This assumption is closely related to the greater evidence for vowels spreading features to preceding consonants than for vowels being affected by preceding consonants in early phonological systems (Ingram, 1989). Therefore, it is expected that CV association will diminish after place features are acquired for most consonants, that is, at around age four.

Based on the theories that make explicit predictions, the following CV affiliation patterns can be described:

(1) Alveolar consonants tend to be followed by front vowels. (However, Kawasaki-Fukumori's acoustically based predictions (1992) claim that this association is disfavoured in the world's languages.)

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

(2) Velar consonants tend to be followed by back vowels. (Both Lindblom's (1990) and Kawasaki-Fukumori's models predict that the association of velars and high back vowel is disfavoured.)
(3) There are no convergent predictions of the association of labial consonants.

Although the underlying mechanisms for neuromuscular control in speech development are unknown, Browman & Goldstein's articulatory phonology, which incorporates various complementary perspectives from different theories, seems to offer the most complete account of the developmental process of CV association. Despite its strengths, this model does not adequately incorporate the crucial role that phonological patterns from ambient languages play in the development of CV affiliation. Establishing a theoretical framework for CV affiliation in speech development will require incorporating models of feature geometries derived from observing cross-linguistic phonological patterns to articulatory phonology, or bringing these two theories closer together and discussing constraints in optimality theory. Furthermore, as suggested by Kingston & Cohen (1992), the validity of this theoretical framework will need to be reassessed in light of future findings on neuromuscular control of speech development and this low-level task dynamic model will need to be connected with high-level control (e.g. control from the central nervous system or from cognitive mechanisms) to establish a comprehensive theory of speech development.

Empirical studies

The earliest discussion of CV associations is found in Jakobson's (1941/68) citation of French and Czech examples. In these examples, labials are found only with back vowels /u, o, ɑ/ (e.g. /papa/), whereas dentals tend to co-occur with front vowels /i, e/ (e.g. /dédé/ Also, the adult model with labial + front vowel sequences is replaced by dental + front vowel sequences (e.g. /dé/ for /dé/) because of children's CV co-occurrence constraints. These regular distribution patterns of dental + front vowels and labial + nonfront vowels have generally been supported by data from English-learning children in Fudge (1969), Braine (1974), and Stoel-Gammon (1983). In addition, the central vowel /ʌ/ was found to follow either labial consonants (e.g. /bʌbʌbi/ in Braine, 1974; and /bʌbu/ for 'bubble' in Stoel-Gammon, 1983) or velar consonants (e.g. /kʌk/ for 'cake' and /gʌŋ/ for 'garden' in Fudge, 1969). Regarding the underlying mechanism for restricted CV sequences, Jakobson's (1941/68:49) explanation of articulatory constraints whereby 'Many children are plainly not able to produce a labial sound before a front vowel' seems to be the first comment on CV co-occurrence in early speech. Later, Stoel-Gammon (1983) suggested that the affiliation of

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

consonants and vowels in early syllables is due to the conditioning effect of vowels on preceding consonants before children's complete mastery of labial and dental contrasts (before 1;5). Nevertheless, these findings and proposals for theoretical bases are preliminary, as most are based on very limited data (fewer than 20 lexical items in each study) from on-site transcriptions of a single subject's production during a short period (1–4 months) of the second year of life.

About two decades after these preliminary studies, CV association recently received attention in the literature. Davis & MacNeilage (1990) conducted the first systematic study of CV association with data from perceptual transcriptions of 13 485 utterances of a single child between 1;2 and 1;8. Based on these data, they generalized three major CV association patterns in both early words and concurrent babbling: (1) labials + central vowels /ʌ, ə, a/; (2) alveolars + high front vowels /i, ɪ/; and (3) velars + high back vowels /u, ʊ/. Davis & MacNeilage further indicated that association (1) is a strong tendency found in monosyllables (V1 and V2 of bisyllabic productions), and that association (3) is weak since it appeared in only a few velars (17% of which were only found in a single lexical item, 'cookie'). Citing these findings, Davis & MacNeilage proposed that the association patterns are based on syllabic frames produced by mandibular oscillation. That is, labials + central vowels represent a pure frame, alveolars + high front vowels constitute a front frame, and velars + high back vowels form a back frame. These association patterns were further confirmed in two consecutive studies: a longitudinal study with 14 431 syllables from six infants (ages 0;6 to 1;0) (Davis & MacNeilage, 1995) and a study of 10 infants during the 50-word stage (Davis, MacNeilage, & Matyear, 2002). These studies came to conclusions similar to those of preceding studies on the association of alveolars + front vowels. However, the strong tendency of labials + central vowels reported by Davis & MacNeilage was not prominent in the findings of any of the four other studies reviewed above.

The association between dentals and front vowels was also supported by Langsdale's (1993) longitudinal and cross-sectional study with nine infants ranging in age from 1;6 to 2;0. However, Langsdale's results on the association between labials and either front or back vowels seem to contradict Davis & MacNeilage's notion of 'pure frame.' In addition, none of the three association patterns in Davis & MacNeilage (1990) were found in a study of CV coarticulation with 387 CV utterances of a single child at 1;0 and 1;9 (Sussman, Minifie, Buder, Stoel-Gammon & Smith, 1996). Sussman *et al.* stated that: (1) labials are found before either high-front or high-back vowels; (2) dentals can co-occur with almost all kinds of vowels (no specific tendency); and (3) velars are followed by high-front or low-back vowels. In sum, although CV association patterns have been reported in

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

several studies with English-learning children, no general agreement has been achieved in the limited database.

After Davis & MacNeilage's (1990) proposed three CV association patterns, two crosslinguistic studies (Vihman, 1992; de Boysson-Bardies, 1993) of CV utterances of children from English, French, Japanese, Swedish, and Yoruban environments sought universal and language-specific patterns of CV associations. The three CV associations suggested by Davis & MacNeilage (1990) were only partially supported by Vihman's (1992) analysis of practiced syllables of 23 infants at 0;9, 0;10, and 0;11 from four language environments. Vihman confirmed Davis & MacNeilage's report of a strong association between labials and central vowels (pure frame) and a weak association between velars and back vowels (back frame). However, Vihman's study did not show a clear tendency for the association between alveolars and front vowels (front frame). Nevertheless, a new association pattern, /h/ with central vowel, was observed in Vihman's study. In de Boysson-Bardies's (1993) crosslinguistic studies with 23 infants between ages 0;10 to 1;0 from four linguistic environments, none of Davis & MacNeilage's CV association patterns were found to generally apply. De Boysson-Bardies found that: (1) labials and central vowels are associated only in French infants, whereas two other patterns of association reflect the target language (labial and front vowels for American infants, labials and back vowels for Swedish infants); (2) a general tendency toward the association of dentals and front vowels in French, American and Swedish infants, but two other patterns of association that reflect target languages (dentals and back vowels, dentals and central vowels) are frequently produced in the second syllable by French, Swedish and Yoruban infants; and (3) velars and back vowels show association patterns, but in languages with more velars, two other patterns are found (velars and central vowels for American and Swedish infants, velars and front vowels for Yoruban infants). In de Boysson-Bardies's study, children's CV utterances already displayed language-specific CV association patterns even before age 1;0. The highly diverse language-specific patterns exhibited in this cross-linguistic study make it impossible to identify any universal CV association patterns. Although two association patterns predicted by Davis & MacNeilage – labial and central vowels, dentals and front vowels – are found in some language groups, they are far from being universal patterns.

Regarding the theoretical basis for CV affiliation in early syllables, the associations between dentals and front vowels and between velars and back vowels observed in English-learning infants and in some crosslinguistic data seem to agree with Clements's (1991) prediction of unified characteristics of the place features of consonants and vowels. In addition, the same-place articulation of CV segments in early syllables and the overcoming of CV co-occurrence constraints in the developmental process support Browman

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

& Goldstein's (1989, 1992) concept of the differentiation and coordination of articulatory gestures. Both these theories assume that early syllables exhibit universal patterns of CV co-occurrence, which are governed by either phonological patterning or gestural control, and that language-specific CV association patterns that are free from universal constraint are acquired only later in development. However, except for the prominent association patterns found in several studies – dentals and front vowels – no systematic universal patterns of CV association can be identified from the diverse conclusions of limited data in previous studies.

A significant limitation in previous studies on CV association is that most were based on Indo-European languages. Before confident conclusions on universality can be drawn, it is important to evaluate the above issues with crosslinguistic data, especially with languages that have substantially different phonological systems from those of Indo-European languages. For example, in contrast to the high frequency of front vowels and labial consonants in English (Vihman, Kay, de Boysson-Bardies, Durand & Sundberg, 1994), Mandarin exhibits overwhelmingly monosyllabic structures, and a high frequency of back vowels and dental consonants (Cheng, 1982). The present study provides empirical data on the acquisition of CV syllables by Mandarin-learning infants to shed light on the issue of CV co-occurrence. In this cross-sectional study, data from 24 infants aged 0;7 to 1;6 were investigated to determine:

(A) if co-occurrence patterns of consonants and vowels exist in the vocalizations of Mandarin-learning infants, and
(B) if such CV co-occurrence patterns exist in Mandarin-learning infants, are they identical to those found in other language groups (and thus based on underlying universal constraints) or do these patterns reflect the patterns of their major linguistic input?

METHODS

Participants

Twenty-four infants (twelve aged 0;7–1;0 and twelve aged 1;1–1;6) were solicited as subjects through informal referral from the surrounding community in Tainan City (Taiwan). Tainan is the fourth-largest city in Taiwan, with a population of 730 000. Letters about the study's purpose and procedure were distributed through friends and neighbours. Interested parents then contacted the investigator for an interview and signed consent forms to participate. All participating infants were from middle-class families. With respect to the infants' language environment, the speech of the caregivers in caregiver–child interactions was also analysed to study the phonetic characteristics of the input language. Although it is well known

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 1. *Subject characteristics* (N = 24)

Subject	Gender	Age (year, month, day)	Number of CV syllables	CV syllables/all syllables	Vocabulary produced (words)
S1	M	0;6.28	37	0·2	1
S2	M	0;7.22	73	0·2	1
S3	F	0;9.07	117	0·5	11
S4	F	0;9.18	66	0·2	5
S5	M	0;9.22	129	0·6	10
S6	M	0;10.06	73	0·4	0
S7	F	0;11.00	119	0·5	3
S8	F	0;11.00	163	0·6	3
S9	M	0;11.18	279	0·9	13
S10	F	1;0.00	137	0·7	9
S11	M	1;0.20	262	0·8	0
S12	M	1;0.25	42	0·2	2
S13	F	1;2.08	141	0·5	60
S14	M	1;3.19	204	0·6	59
S15	M	1;3.24	309	0·7	65
S16	F	1;4.04	224	0·6	47
S17	M	1;4.19	88	0·5	11
S18	F	1;4.23	179	0·6	60
S19	F	1;5.00	438	0·7	76
S20	F	1;5.02	649	0·7	73
S21	M	1;5.11	590	0·8	62
S22	F	1;5.15	354	0·8	87
S23	M	1;5.18	169	0·5	5
S24	M	1;5.23	149	0·3	16

that the nature of speech directed to infants differs from that directed to adults (especially in prosodic features), infant-directed speech was analysed in this study to provide information on the patterns in the ambient language used by adults. The characteristics of infant-directed speech may interact with the properties of the ambient language, but it seems that these characteristics, which affect mainly prosodic features, would not materially affect CV co-occurrence patterns that were the main issue in this study. Subject information is listed in Table 1.

The rationale for choosing this age range (0;7–1;6) was that infants begin to produce speech-like CV sequences around 0;6–0;7 and they overcome CV co-occurrence constraints and produce their first 50 words by 1;6. The range of 0;7–1;6 is thus appropriate to capture the developing vocalizations of CV sequences. The division of the age range (six months before and six months after age 1;0) was chosen to observe developmental changes in infants' vocalizations for a separate paper. The size of each age group was felt to be adequate to exclude data reflecting extreme individual differences in phonetic development.

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

The following criteria were used to exclude extreme individual differences among participating infants. First, the infants had normal hearing and were free of any developmental delays, with no clinical history of remarkable postnatal illness, as determined through parental interviews and observation of children during data collection. According to caregivers' general observations, the infants were also healthy and in a stable emotional state before and during the days on which recordings were made. Second, to have valid comparisons of the value of proportions from different infants, the number of CV syllables produced by each infant was calculated first as a basic criterion. Based on Vihman's (1992) crosslinguistic study, children in the first age group (0;7–1;0) who produced fewer than 25–30 CV syllables in a 45-minute recording were excluded from the study. Also excluded were children in the second age group (1;1–1;6) who produced fewer than 40 CV syllables in a 45-minute recording.

The third criterion used to eliminate extreme individual differences was an age-appropriate canonical babbling ratio (the ratio of canonical syllables to total number of syllables), based on averages. Based on previous studies (Lewedag, Oller & Lynch, 1994), infants at ages 0;7–1;0 with a canonical babbling ratio of less than 0·2 were excluded from the analysis, as were infants at ages 1;1–1;6 with a canonical babbling ratio of less than 0·3. The fourth criterion was lexical development based on parental reports. A translated and modified version of the MacArthur Communicative Development Inventory: Infants scale (Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly, 1993) was used as an index of the developmental process. This inventory was modified by the investigator to adjust for language and cultural differences between English and Mandarin. In addition, caregivers were encouraged to include infants' words that were not included in this scale. According to this scale, the average number of words produced by infants ages 0;7–1;0 and infants ages 1;1–1;6 are about 0 and 1–25 words, respectively. All subjects recruited for this study in the two age groups met these criteria. The fifth criterion pertains to the linguistic environment. To exclude substantial confounding factors in language input, this study included only infants whose primary caregivers spoke mainly Mandarin, especially when interacting with the infants. Those infants whose primary caregivers mainly used Taiwanese when interacting with infants were excluded.

Data collection

Infants' vocalizations were audio recorded during observations of their natural daily activities in their homes or day-care centres. On average, a 45-minute recording was made from two-hour observation sessions for each of the 24 infants. The choice of home environments and day-care centres as

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

observation settings is based on a report (Lewedag *et al.*, 1994) that infants vocalize more than twice as much and produce well-formed syllables more frequently in home environments when familiar adults are present than in lab settings. Recordings were made in a quiet room in the home (either the living room or bedroom) or day-care centre, with only caregivers, the infant and the investigator present, though sometimes older children were around. Observation schedules were developed according to caregivers' suggestions regarding the best times for infants to vocalize. However, observations were often rescheduled, since the infants were not awake or in a suitable state. In such cases, second and third attempts were made within a two-week period.

With respect to the duration of recording sessions, two-hour (Davis & MacNeilage, 1990), one-hour (Kent & Bauer, 1985), and half-hour sessions (Vihman, Macken, Miller, Simmons & Miller, 1985; Vihman, 1992) have all been used in previous longitudinal or cross-sectional studies. In cross-sectional research, like the present study, longer times are required for each session to obtain sufficient data because each observation is a new situation for infants, infant's family and the investigator. Choosing good recording conditions, instructing caregivers, and becoming familiar with the patterns of infant's vocalizations requires more than one hour of total visitation time. Therefore, two-hour observation sessions were scheduled. A 45-minute audio recording was made for each child in each age group based on experience gained from a pilot study, in which approximately 45 minutes represented the maximum amount of data that could be obtained in a two-hour observation session for infants at about age 1;0. Hodge (1989) also found 45-minute recordings to be adequate for providing a representative sample of an infant's vocalizations in the home environment.

No specific or systematic planning of elicitation was done in this study, except for natural elicitations in daily life (e.g. naming toys or pictures in the infants' home environments). The goal was to ensure that spontaneous vocalizations, as opposed to elicited and imitated sounds, were recorded. According to Delack (1976) and our pilot study, infants tend to babble more when they are alone or when there is extended silence in an interaction. Therefore, before recording, caregivers were instructed to allow natural pauses and silent moments in the interaction. To provide necessary supplementary information for later transcription, the infants' repeated vocalizations of the same syllables were transcribed on-site to capture the first-hand production (Ingram, 1989) whenever possible. Such on-site transcriptions then served as a reference for later transcriptions from the audio recordings. In addition, nonlinguistic contexts that might help identify the infants' attempted forms were noted on-site.

To achieve good quality and flexible recording without distracting subjects, a wireless omni-directional lapel microphone (Telex ProStar

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

Wireless System R-10) was clipped to a bib, approximately six inches from the mouth, with the transmitter hidden in the pocket of the bib. The microphone was linked to a Sony digital audio tape-corder (DAT recorder TCD-D8). The DAT recorder was used for providing high recording quality for perceptual transcription in this study and future acoustic analysis. Since adult speech is easy to recognize for transcription, and since caregivers were always around the infants during the observation sessions, child-directed speech was recorded through the same omni-directional microphone used to record infants' vocalizations. In addition, the recorder was always powered with batteries rather than a cable, to provide more flexibility in choosing observation locations in the rooms and in accommodating the environments of different homes or day-care centres.

Data analysis

For analysis of the frequency of the various patterns of CV co-occurrence, both infants' vocalizations and adults' speech were transcribed and organized with respect to six criteria used in data selection. First, to incorporate all possible precursors of later speech in infants' vocalizations, an inclusive definition of 'speech-like' sounds was adopted in selecting data for analysis. This broad definition excludes only vegetative or reflexive sounds (e.g. cries, coughs, breathing noise, etc). Second, among all speech-like vocalizations, only sequences with at least one C-like and one V-like sound (canonical syllables) were analysed. In identifying canonical syllables in infants' vocalizations, perceptual judgments about well-formed syllables were made based on general criteria used in previous studies (Davis & MacNeilage, 1990; Vihman, 1992). That is, isolated vowels and sequences of vowels preceded or followed by glottal stops were not included in the database. Third, all spontaneously produced canonical syllables were transcribed and analysed, whereas imitated data were not included. Fourth, each of the CV syllables in isolated syllables, reduplicated babbles or variegated babbles were analysed as individual CV syllables. Fifth, no distinction between babbling and the production of early words was noted in analysing data, and the infants' CV productions were not analysed in terms of accuracy with reference to attempted adult words. Sixth, all spontaneous productions of CV sequences during the 45-minute recordings of each of 24 infants were phonetically transcribed. For those recordings longer than 45 minutes, the 45-minute section with the most CV syllables was chosen for transcription and analysis.

With regard to phonetic transcription, broad transcription was used, based on the IPA system. To allow comparisons with previous studies (e.g. Davis & MacNeilage, 1990; Vihman, 1992), special attention was given to place contrasts among consonants and vowels (labials, alveolars, velars

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 2. *Inter-transcriber reliability and intra-transcriber reliabilities*

	Inter-transcriber reliability with 1052 CV syllables	Intra-transcriber reliability with 1245 CV syllables
Consonant place of articulation (labial, alveolar, velar)	98·8%	99·8%
Consonant manner of articulation (stop, nasal, fricative, affricate, lateral)	97·5%	99·0%
Vowel backness (front, central, back)	78·1%	95·4%
Vowel height (high, mid, low)	71·6%	95·5%
Intonation pattern (6 intonation categories)	—	92·0%

for consonants; front, central, back for vowels), and transcriptions of consonants or vowels that presented ambiguity in place of articulation were checked again or deleted from the sample. Softly articulated CV syllables were also excluded from analysis because these syllables could not be clearly identified perceptually without referring to vocal gestures. Moreover, strings of perceptually unidentifiable consonants and vowels were excluded. The percentages of discarded data ranged from 0 to 34%. These transcription criteria were applied throughout the study and in retranscription, for intra-transcriber reliability. To estimate intra-transcriber reliability, the investigator checked the consistency of the perceptually-based transcriptions by retranscribing a random five-minute sample (11% of the total recordings) for each of the 24 infants. In addition, to check inter-transcriber reliability in the transcription of the major categories of consonants and vowels (similar to the criteria used in Vihman *et al.*, 1985) and to check for possible language bias, a second experienced transcriber with a different linguistic background (English) transcribed a randomly selected five-minute sample (11% of the total) from each infant's data. The second transcriber had no experience with the Mandarin phonological system and no knowledge of this study's purpose. Both the intra-transcriber and the inter-transcriber reliabilities are listed in Table 2.

The high degree of agreement between transcribers in identifying the low vowel /a/ and affricates is worth noting because the prominence of these two categories in Mandarin-learning infants is one of the major findings. For analysis of child-directed speech, the caregivers' speech in a 45-minute recording was transcribed based on the IPA system. The total number of CV syllables transcribed for 24 infants and 24 caregivers is represented in Table 3. The number of CV syllables produced by the 24 infants of this study was positively correlated with that of their caretakers, as indicated by a one-tail Spearman correlation ($r = 0{\cdot}614$, $p = 0{\cdot}001$).

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 3. *Numbers of CV syllables produced by 24 infants and 24 caregivers*

	Infants	Caregivers
S1	37	167
S2	73	294
S3	117	798
S4	66	430
S5	129	1550
S6	73	471
S7	119	618
S8	163	86
S9	279	1273
S10	137	1886
S11	262	600
S12	42	880
S13	141	669
S14	204	1822
S15	309	1512
S16	224	983
S17	88	130
S18	179	727
S19	438	2963
S20	649	1686
S21	590	965
S22	354	1086
S23	169	868
S24	149	944

To answer research question (A) (Do co-occurrence patterns of consonants and vowels exist in the vocalizations of Mandarin-learning infants?), CV sequences were arranged into nine cells (Table 4). The percentages for the nine possible CV combinations out of the total observations of CV sequences were calculated.

Because observations of CV co-occurrence patterns among the nine cells were not independent, chi-square or ANOVA procedures (used in previous studies) are not appropriate for this research question. Instead, one-sample Wilcoxon tests were used to compare percentages of co-occurrence in the three cells of each row, to detect any CV co-occurrence patterns. The rationale for choosing a nonparametric test is as follows. First, the sample size in this study is not large enough to justify assuming a normal distribution of the population as for most parametric tests (including two-sample *t* tests and ANOVA *F* tests, Marascuilo & Serlin, 1988: 482). Second, the profound individual differences expected among young infants violate the assumption of equality of population variances (homogeneity of variance) for most parametric tests. Therefore, as recommended by Marascuilo & Serlin (1988:294), nonparametric tests that are distribution-free and have minimal assumptions were used in this study. To answer question (B) (Are the

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 4. *Nine consonant–vowel co-occurrence patterns*

	front vowel	central vowel	back vowel
labial	labial + front	labial + central	labial + back
alveolar	alveolar + front	alveolar + central	alveolar + back
velar	velar + front	velar + central	velar + back

patterns of CV co-occurrence found in this study identical with those found in other language groups?), if the CV co-occurrences in Mandarin-learning infants exhibited patterns similar to those reported by Davis & MacNeilage (1990), then cells (labial + central vowel), (alveolar + front vowel), and (velar + back vowel) in Table 4 would be expected to show the greatest medians among the three cells in the same rows. Finally, to determine whether CV co-occurrence found in Mandarin-learning infants reflected their input language patterns, the CV sequences in child-directed speech were examined first, following the same procedure as described above for the infant groups. Two-sample Wilcoxon tests were then used to compare percentages of these CV co-occurrence patterns with those from infants' data.

Mandarin consonant–vowel co-occurrences

There are strict consonant–vowel co-occurrence patterns in Mandarin. According to traditional Chinese phonology (Luo, 1992), patterns of CV co-occurrence are based on place of articulation of consonants and the backness and rounding of the immediately succeeding vowel (either a nucleus or a medial). The following co-occurrence patterns in CV syllables are based on tables provided by Luo (1992). Consonants are grouped into three major categories for ease of comparison with infants' vocalizations: labials, alveolars and velars. 'Alveolars' are used as a general category that includes dentals, post-alveolar retroflexes, and palatals. Since the infant subjects did not have a full set of dentition, the three pairs of unaspirated/aspirated affricates in adults' system are grouped as one category (alveolar or coronal) due to sharing one property – raising the tongue above its presumed rest or neutral position. Moreover, 'alveolar' is used for ease of comparison with previous studies. In addition, to correspond to analysis of the infants' data, vowels are also grouped into three major categories: front, central, and back. The vowel /a/, the only low vowel in Mandarin, is regarded as a low-back vowel to differentiate it from the resting position of the central vowel /ə/, which occurs frequently in infants' data.

(1) Among alveolars, the palatal affricates and fricative /tɕ, tɕ', ɕ/are followed by /i/ or /y/, whereas the post-alveolar retroflexes /tʂ, tʂ', ʂ/ and dentals /ts, ts', s/ never occur before /i/, or /y/.

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

(2) Other alveolars /t, t', n, l/ can be followed by either back vowels /ɤ, u, ɑ/ or front vowels /i, y/. Therefore, alveolars as a category do not display a specific co-occurring pattern.
(3) Labials /p, p', m, f/ are followed by either back vowels /u, ɑ, o/ or the front vowel /i/. With respect to the co-occurrence of labial features, although labials can co-occur with round vowels /u, o/, they are never found preceding the high-front round vowel /y/.
(4) Velars /k, k', x/ are found only before back vowels /ɤ, u, ɑ/, and are never followed by front vowels /i, y/.

Reanalysis of Cheng's (1982) data shows patterns similar to those described by Luo (1992) in frequency of occurrence:

(1) Alveolars are equally distributed in front and back vowel contexts (alveolars with front vowels 21%, alveolars with back vowels 35·8%).
(2) Labials show no specific co-occurrence patterns with either front vowels (4·6%) or back vowels (7%).
(3) Velars show a greater tendency to co-occur with back vowels (9·4%) than with front vowels (0·1%).

In sum, alveolars can co-occur in Mandarin with either front or back vowels, and both labials and velars display a pattern of co-occurring with back vowels.

RESULTS AND DISCUSSION

CV co-occurrence patterns were observed in both the infant and adult data. In particular, labials tended to co-occur with back vowels, and alveolars were frequently followed by front vowels. The less frequently occurring consonants (i.e. velars), when they were produced, tended to co-occur with back vowels. As shown in Table 5, the sequence of labials + back vowels (median = 22·1% in infants; median = 14·1% in adults) occurred more frequently than the sequences of labials + front vowels or labials + central vowels in both the infants' and caregivers' data.

Alveolars occurred in the same order of frequency in both infants' and adults' data. That is, the sequence of alveolars + front vowels (median = 25·1% in infants; median = 26% in caregivers) was produced more frequently than the sequence of alveolars + back vowels (median = 9% in infants; median = 21% in caregivers). In addition, the pattern of alveolars + central vowels was produced least frequently by both infants and adults. Regarding velars, the patterns of velars + back vowels and velars + central vowels occurred more frequently than the pattern of velars + front vowels in both the infants' and caregivers' data. These differences were all significant by one-sample Wilcoxon tests ($p < 0·05$).

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 5. *Median differences in nine consonant–vowel co-occurrence patterns produced by infants and caregivers* (N = 48)

Consonant–vowel co-occurrence patterns	Infants Median (%)	Caregivers Median (%)	t_w*
labial + front	0·75	3·45	444·5†
labial + central	1·35	1·70	556·0
labial + back	22·10	14·10	697·5
alveolar + front	25·05	25·95	574·0
alveolar + central	2·15	9·00	394·0†
alveolar + back	8·95	21·00	441·5†
velar + front	0·00	1·80	413·5†
velar + central	0·65	2·90	422·5†
velar + back	2·30	14·25	408·0†

* Two-sample Wilcoxon test, nondirectional.
† $p < 0{\cdot}006$ (0·05/9).

The most frequently occurring vowels in the infants' data were /ɛ, ə, ɑ/. In the co-occurrence of labials + back vowels, the majority of back vowels in the infants' data involved the vowel /ɑ/. In the caregivers' data, the co-occurrence of labials and low back vowels also occurred more frequently than the combination of either labials with high back vowels or labials with mid back vowels. Nevertheless, there are several differences between infants' data and adults' child-directed speech. In the infants' data, the central vowel /ə/ occurred second-most frequently in combination with labials, whereas in the caregivers' data, the high back vowel /u/ occurred second-most frequently in combination with labials (after low-back vowels), with no obvious tendency in the co-occurrence of labials and mid-central vowels, as was found in the infants' data. In addition, the co-occurrence of alveolars and front vowels differed markedly between the infants' and caregivers' data. Infants produced many more front vowels /ɛ/ than front vowels /i/ in combination with alveolars. Conversely, adults used high-front vowels much more frequently than mid-front vowels in combination with alveolars. In fact, in child-directed speech the co-occurrence pattern of alveolars + high-front vowels occurred most frequently among all CV co-occurrence patterns, whereas the combination of alveolars and mid-front vowels was among the least frequently occurring patterns. With respect to velars, although both infants and caregivers more frequently produced the back vowel /ɑ/ and the central vowel /ə/ in combination with velars, the infants' and caregivers' data displayed different co-occurrence patterns, which were parallel to the patterns of labials + vowels as described above. That is, in the infants' data, back vowels occurred most frequently when produced with velars, followed by the central vowel, whereas in the caregivers' data, high-back vowels (instead of mid-central vowels) occurred second-most

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 6. *Mean consonant–vowel co-occurrence productions by nine major patterns and their correlation between infants and caregivers* (N = 48)

Consonant–vowel co-occurrence patterns	Infants (%)*		Caregivers (%)*			
	Mean	S.D.	Mean	S.D.	*r*†	*p*
labial + front	30·8	5·28	4·10	2·09	0·26	0·114
labial + central	7·05	17·69	1·93	1·32	−0·06	0·382
labial + back	32·58	29·37	15·68	7·77	0·46	0·013
alveolar + front	25·58	21·09	25·65	4·69	−0·01	0·480
alveolar + central	5·74	10·66	10·42	4·27	0·26	0·114
alveolar + back	13·40	13·77	21·77	5·12	−0·42	0·022
velar + front	1·21	3·33	1·83	1·17	−0·11	0·307
velar + central	3·18	7·33	4·74	4·54	0·22	0·155
velar + back	7·46	13·25	13·88	5·05	0·34	0·051

* The sum of the mean percentages within each column is 100.
† Spearman *r*, directional.

frequently (after low-back vowels) in combination with velars. This discrepancy between the infants' and caregivers' data was also found in CV co-occurrence patterns involving the velar fricative /x/.

Strong positive correlations were found between the infants' and caregivers' data for labials + back vowels and velars + back vowels (Table 6). Also, two-sample Wilcoxon tests (Table 5) indicated no significant differences between the infants' and caregivers' data in combinations of labials + back vowels, labials + central vowels, and alveolars + front vowels. In sum, despite the differences described above, the infants' and caregivers' data displayed similar patterns of CV co-occurrence, especially in the co-occurrence of alveolars + front vowels, labials + back vowels, and velars + back vowels.

Alveolars and front vowels

Some CV co-occurrence patterns may be natural outcomes of the frequency distribution of consonants and vowels in a particular language. For example, Mandarin-learning infants in this study produced more back than front vowels (ratio of back to front vowels = 1·7). This pattern of frequency distribution of vowels was observed in the environments of both labials and velars. However, instead of following this general pattern of vowel frequency distribution regardless of consonant environment, infants produced many more front vowels than back vowels in combination with alveolars. The ratio of alveolars + front vowels to alveolars + back vowels was 1·9. In fact, the front vowels produced by infants in this study were found almost exclusively in combination with alveolars (about 86%) in CV syllables.

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 7. *Comparison of consonant–vowel co-occurrence productions by nine major patterns in two infant age groups* (N = 24)

Consonant–vowel co-occurrence patterns	G1 infants (ages 0;7 to 1;0) Median (%)	G2 infants (ages 1;1 to 1;6) Median (%)	t_w*
labial + front	0·00	3·50	111·5
labial + central	4·50	0·45	180·0
labial + back	41·20	19·30	177·5
alveolar + front	15·00	27·85	126·0
alveolar + central	0·20	2·75	142·0
alveolar + back	5·45	14·85	126·0
velar + front	0·00	0·60	117·0
velar + central	0·00	0·90	118·5
velar + Back	0·00	8·10	98·0†

* Two-sample Wilcoxon test, nondirectional.
† $p < 0{\cdot}006$ (0·05/9).

This finding is important evidence for a consonant–vowel association. As shown in Table 7, the predominance of front vowels over back vowels in combination with alveolars is evident in the data for both the younger (aged 0;7–1;0) and older (aged 1;1–1;6) infants. This pattern is especially obvious in the older group (in which the pattern occurred most frequently of all CV co-occurrence patterns) and is closely related to the prominence of alveolar and front vowel combinations in the ambient language. This co-occurrence pattern was very prominent in both the group and individual data for each infant studied, possibly because this association pattern results from both a natural tendency and the pattern of the input language.

The combination of alveolars and front vowels appears to be a universal (or nearly universal) CV association pattern that is supported by various empirical studies (Davis & MacNeilage, 1990; Langsdale, 1993) and is consistent with a number of theoretical proposals (Janson, 1986; Davis & MacNeilage, 1990; Clements, 1991; Davis & MacNeilage, 1995). However, the combination of alveolars + front vowels does not apply to all language groups. For example, Sussman *et al.* (1996) found that alveolars co-occur with front, central, or back vowels (i.e. no specific prominent pattern was identified). Moreover, no evident association of alveolars + front vowels and no universal patterns of CV association involving alveolars found in two crosslinguistic studies on infants from English, French, Japanese, Swedish, and Yoruban environments (Vihman, 1992; de Boysson-Bardies, 1993). These contradictory findings reflect both Kawasaki-Fukumori's (1992) acoustically based CV affiliation framework and Lindblom's model (1992) of low articulatory cost and high discriminability. The former suggests the combination of alveolars + front vowels is generally disfavoured in the world's

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

languages due to the small magnitude of spectral changes, and the latter predicts that alveolars tend to co-occur with either front or back vowels.

Labials and back vowels

Our study found a clear tendency toward combining labials and back vowels in the infants' data. The co-occurrence of labials and central vowels (mostly mid-central vowels) mainly occurred in infants younger than 1;0 and only occurred second-most frequently in infants aged 0;7–1;0 (after the combination of labials and back vowels) (Table 7). The frequency of this combination became less obvious at 1;1–1;6, thus matching the pattern observed in the child-directed speech, in which there was no obvious tendency for labials and central vowels to co-occur.

The combination of labials and back vowels was also found in Jakobson's (1941/68) examples of French and Czech children. This association pattern agrees with Janson's (1986) articulatory constraints on CV sequences: the association of labials and back vowels is favoured. However, our findings are not consistent with the predictions of several other theories. First, study findings are in direct contrast to Kawasaki-Fukumori's (1992) framework of acoustically based CV affiliation, which proposes that the combination of labials and back vowels is disfavoured because of the low levels of spectral change. Second, our findings disclose a tendency toward the co-occurrence of labials with low-back vowels instead of with rounded vowels (high- or mid-back vowels /u, o, ɔ/, as found in Clements's (1991) feature geometry. Third, our finding of the predominance of back vowels over front vowels in combination with labials contradicts Lindblom's (1992) prediction of CV associations according to the model of articulatory simplicity and distinctiveness. This model proposes that labials tend to co-occur with either front or back vowels. Fourth, the combination of labials and back vowels is not predicted by Davis & MacNeilage's frame-then-content theory. Instead, that theory suggests the co-occurrence of labials and central vowels (pure frame) is frequently found in the early phonetic system. Both Davis & MacNeilage's (1990) study of English-learning infants and Vihman's (1992) crosslinguistic study indicate a strong tendency for labials and central vowels to associate.

Based on the theoretical proposals and empirical evidence cited above, we suggest a developmental process for the CV sequences involving labials. There is no single convergent association pattern in labials across languages, as found in de Boysson-Bardies (1993), because language-specific patterns exert an early developmental influence. In data from Mandarin-learning infants, the prominence of the default pattern (labials + central vowels) is gradually overridden by the predominance of the co-occurrence of labials and back vowels occurring in the ambient language. This developmental

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

change occurs before 1;0 and is similar to the CV co-occurrence patterns with velars as discussed in the next section.

Velars and back vowels

Velars occurred much less frequently than labials and alveolars in infants' vocalizations. The only prominent CV association pattern we found for velars was the co-occurrence of velars and back vowels. As seen in Table 7, this co-occurrence pattern did not emerge until infants produced sufficient velars after age 1;0. Velars co-occurred with central vowels (mainly the mid-central vowel) in infants younger than 1;0. However, after 1;0, following the emergence of the combination of velars and back vowels, the combination of velars and central vowels became the second most frequently occurring pattern, revealing a much lower frequency than the combination of velars and back vowels. Based on the empirical evidence above, in Mandarin-learning infants' data, the default association pattern of velars + central vowels, like the CV association of labials, is gradually overridden by the pattern of velars + back vowels in the developmental process, influenced by the prominence of this pattern in the ambient language. The default pattern of velars and central vowels found in younger infants was also found in the CV syllables involving the velar fricative /x/. The velar fricative /x/ is developed late and occurs much less frequently than velar stops. The association pattern of /x/ + central vowels was only found in infants older than 1;0. The late development of /x/ and the CV association involving /x/ reveal a default CV association pattern that does not reflect influence from the pattern in the ambient language (i.e. /x/ + the low back vowel, and /x/ + the high back vowel in the child-directed speech). This default association tendency is similar to the finding in Vihman's (1992) crosslinguistic study. Although our findings offer only limited evidence of this co-occurrence pattern, it might indicate a universal pattern. Further empirical studies with more /x/ productions would be necessary to verify this finding.

The association pattern of velars + back vowels was predicted by several theoretical models (Janson, 1986; Davis & MacNeilage, 1990, 1995; Clements, 1991). However, the combination of velars + back vowels is not the only prominent CV co-occurrence pattern involving these consonants. Velars are also found with front and central vowels in studies of different language groups. For example, velars were followed by either high-front or low-back vowels for an English-learning infant in Sussmam *et al.* (1996) and for Yoruban infants in Vihman (1992). Moreover, the default co-occurrence pattern mentioned above, velars + central vowels, is also found, for example, in English-learning infants (Fudge, 1969; Langsdale, 1993) and in both English-learning and Swedish-learning infants (Vihman, 1992).

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

TABLE 8. *Comparison of our findings with those of previous studies on CV co-occurrence*

Study	labial + back vowel	alveolar + front vowel	velar + back vowel
Frame-then-content (Davis & MacNeilage, 1990, 1995)	X	O	O
Unified features for consonants and vowels (Clements, 1991)	X	O	O
Acoustically-based CV co-occurrence (Kawasaki-Fukumori, 1992)	X	X	—
Articulatory constraints for CV co-occurrence (Janson, 1986)	O	O	O
Low articulatory cost and high discriminability for CV syllables (Lindblom, 1990)	#	#	#
Jakobson (1941/68) [French, Czech]	O	#	—
Fudge (1969) [English]	O	#	X
Braine (1974) [English]	#	O	X
Stoel-Gammon (1983) [English]	#	O	—
Davis & MacNeilage (1990) [English]	X	O	O
Langsdale (1993) [English]	#	O	#
Sussman *et al*. (1996) [English]	#	#	#
Vihman (1992) [crosslinguistic]	X	#	O
de Boysson-Bardies (1993) [crosslinguistic]	#	#	#

Note: O: consistent with our findings.
X: inconsistent with our findings.
#: no specific co-occurrence patterns (distributed patterns).
—: no information.

The above comparison of this study's findings with the predictions of various theoretical models and with the findings of empirical studies on CV co-occurrence is summarized in Table 8.

In summary, the findings in this study offer further empirical evidence for both Davis & MacNeilage's (1990, 1995) frame-then-content theory and Clements's (1991) unified feature framework with regard to the co-occurrence of alveolars + front vowels and velars + back vowels. Both the front frame (alveolars + front vowels) and the back frame (velars + back vowels) in Davis & MacNeilage's theory are supported in the current study. However, only our data for younger infants support the pure frame (labials + central vowels). Our findings are closely related to Clements's predictions of CV co-occurrence of alveolars and front vowels (which share the coronal feature) and of the co-occurrence of velars and back vowels (which share the dorsal feature). However, the predicted combinations of labials and high- or mid-back vowels (rounded vowels) are found only in our data for older infants.

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

These inconsistencies suggest the need to reexamine both the 'pure frame' set forth in Davis & MacNeilage (1990, 1995) and the grouping of labials and rounded vowels posited by Clements (1991) to account for these new linguistic data. Evaluating our findings in light of Janson's (1986) articulatory-based framework and Kawasaki-Fukumori's (1992) acoustical-based framework for CV affiliation in the world's languages, ease of production seems to play a more important role than ease of perception in the development of early CV syllables. All three major CV association patterns suggested in Janson's framework are found in this study; nevertheless, our findings directly contradict the disfavoured CV association patterns indicated in Kawasaki-Fukumori's model. The present study thus suggests that CV co-occurrence patterns in early CV syllables are based on: (1) articulatory constraints (e.g. frame-then-content theory), (2) sensorimotor schemes and gestural overlapping at early stage of development (Browman & Goldstein, 1989, 1992), and (3) the CV affiliation patterns in the world's language determined by articulatory and acoustical bases (e.g. Clements's feature geometry), especially the former (e.g. Janson, 1986).

There is insufficient empirical evidence in our study or in the available literature to determine any truly universal CV co-occurrence pattern. Nevertheless, three co-occurrence patterns found in the early stage of development in this study appear to be default patterns and are very likely universal. These include alveolars + front vowels, labials + central vowels, and velars + central vowels (including the co-occurrence of /x/ and central vowels). However, the high frequency of both labials + central vowels and velars + central vowels is only found in the early stage with a language-specific pattern (labials + back vowels, and velars + back vowels) becoming prominent later in development. Therefore, our findings and those in the literature suggest that universal CV co-occurrence patterns may prevail at a very early stage (0;7–1;0) but only for a short period. However, no single universal pattern seems to persist and to apply in most language environments in the world because the influence of patterns in the ambient language is evident throughout the developmental process, even in early infancy.

As indicated in this study, infants' vocalizations are influenced by both overall productivity (shown by the positive correlation between number of CV productions of infants and their caretakers) and the phonetic content of the ambient language. The early and clear language-specific tendency found in Mandarin-learning infants soon after they start to produce CV syllables might be closely related to Mandarin's unique monosyllabic structure. The mutual reinforcement of early-developed lexical tone perception and production, and the prominence of syllables in perception may facilitate the development of segmental perception in two ways: first, by providing cues for syllable organization in speech, and, second, by creating tone classes in which segmental structures can be recognized. In addition, our finding

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

that clear CV co-occurrence patterns develop early provides valuable information for early detection of phonological disorders by age 2;0. Further studies are suggested to explore the clinical implications of findings concerning the early emergence of both universal and language-specific patterns.

REFERENCES

Braine, M. D. S. (1974). On what might constitute a learnable phonology. *Language* **50**, 270–300.

Browman, C. P. & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology* **6**, 201–51.

Browman, C. P. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica* **49**, 155–80.

Cheng, C.-M. (1982). Analysis of present-day Mandarin. *Journal of Chinese Linguistics* **10**, 281–358.

Clements, G. N. (1991). Place of articulation in consonants and vowels: a unified theory. *Working Papers of the Cornell Phonetics Laboratory* **5**, 77–123.

Davis, B. L. & MacNeilage, P. F. (1990). Acquisition of correct vowel production: a quantitative case study. *Journal of Speech and Hearing Research* **33**, 16–27.

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. E. & 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.

de Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage & J. Morton (eds), *Developmental neurocognition: speech and face processing in the first year of life*. Dordrecht: Kluwer.

Delack, J. (1976). Aspects of infant speech development in the first year of life. *Canadian Journal of Linguistics* **21**, 17–37.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S. & Reilly, J. S. (1993). *MacArthur communicative development inventories*. San Diego, CA: Singular.

Fudge, E. C. (1969). Syllables. *Journal of Linguistics* **5**, 253–86.

Gierut, J. A., Cho, M. & Dinnsen, D. (1993). Geometric accounts of consonant–vowel interactions in developing systems. *Clinical Linguistics and Phonetics* **7**(3), 219–36.

Halle, M. (1992). Phonological features. In W. Bright (ed.), *International encyclopedia of linguistics*, Vol. 3. New York: Oxford University Press.

Hodge, M. (1989). *A comparison of spectral-temporal measures across speaker age: implications for an acoustic characterization of speech maturation*. Unpublished doctoral dissertation, University of Wisconsin-Madison.

Ingram, D. (1989). *Phonological disability in children*. Second edition. The Hague: Mouton.

Jakobson, R. (1941). *Child language, aphasia, and phonological universals*. The Hague: Mouton (English translation 1968).

Janson, T. (1986). Crosslinguistic trends in the frequency of CV sequences. *Phonology Yearbook* **3**, 179–95.

Kawasaki-Fukumori, H. (1992). An acoustical basis for universal phonotactic constraints. *Language and Speech* **35**(1–2), 73–86.

Kent, R. D. & Bauer, H. R. (1985). Vocalizations of one-year-olds. *Journal of Child Language* **12**, 491–526.

Kingston, J. & Cohen, A. H. (1992). Extending articulatory phonology. *Phonetica* **49**, 194–204.

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press

Lahiri, A. & Evers, V. (1991). Palatalization and coronality. In C. Paradis & J. F. Prunet (eds), *Phonetics and phonology, Vol. 2; the special status of coronals: internal and external evidence*. New York: Academic Press.
Langsdale, T. E. (1993). *Vowel production in relationship to consonant context*. Unpublished master's thesis, University of Nevada, Reno.
Lewedag, V. L., Oller, D. K. & Lynch, M. P. (1994). Infants' vocalization patterns across home and laboratory environments. *First Language* **14**, 1(40), 49–65.
Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H & H theory. In W. J. Hardcastle & A. Marchal (eds), *Speech production and speech modeling*. Dordrecht: Kluwer.
Lindblom, B. (1992). Phonological units as adaptive emergent of lexical development. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development: models, research, implications*. Timonium, MD: York Press.
Lindblom, B., MacNeilage, P. & Studdert-Kennedy, M. (1983). Self-organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie & O. Dahl (eds), *Explanations of linguistic universals*. The Hague: Mouton.
Luo, C.-C. (1992). *Kuo Yü hsüeh* (*Studies of Mandarin Chinese*). Taipei: Wu-Nan.
MacNeilage, P. (1996). Formal versus functional approaches to the acquisition of speech production. Paper presented at the Symposium on Research in Child Language Disorders.
MacNeilage, P. (1997). Acquisition of speech. In W. J. Hardcastle & J. Laver (eds), *The handbook of phonetic sciences*. Cambridge, MA: Blackwell.
Marascuilo, L. A. & Serlin, R. C. (1988). *Statistical methods for the social and behavioral sciences*. New York: W. H. Freeman and Company.
Matyear, C. L., MacNeilage, P. F. & Davis, B. L. (1998). Nasalization of vowels in nasal environments in babbling: evidence for frame dominance. *Phonetica* **55**, 1–17.
Stoel-Gammon, C. (1983). Constraints on consonant–vowel sequences in early words. *Journal of Child Language* **10**, 455–57.
Sussman, H. M., Minifie, F. D., Buder, E. H., Stoel-Gammon, C. & Smith, J. (1996). Consonant–vowel interdependencies in babbling and early words: preliminary examination of a locus equation approach. *Journal of Speech and Hearing Research* **39**, 424–33.
Vihman, M. M. (1992). Early syllables and the construction of phonology. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development: models, research, implication*. Timonium, MD: York Press.
Vihman, M. M., Kay, E., de Boysson-Bardies, B., Durand, C. & Sundberg, U. (1994). External sources of individual differences? A crosslinguistic analysis of the phonetics of mothers' speech to one-year-old children. *Developmental Psychology* **30**, 651–62.
Vihman, M. M., Macken, M. A., Miller, R., Simmons, H. & Miller, J. (1985). From babbling to speech: a reassessment of the continuity issue. *Language* **61**, 397–445.

https://doi.org/10.1017/S0305000905006896 Published online by Cambridge University Press