Title
Functional load and frequency predict consonant emergence across five languages

Journal
UC Berkeley PhonLab Annual Report, 13(1)

Author
Cychosz, Margaret

Publication Date
2017

DOI
10.5070/P7131040758

Functional load and frequency predict consonant emergence across five languages

Margaret Cychosz
University of California, Berkeley

Abstract

Frequency can often predict when children will acquire units of language such as words or phones. An additional predictor of speech development may be a phone's functional load (FL), or the contrastive work that a sound performs in a language. A higher FL may correlate with earlier phone emergence in child speech as children selectively converge upon highly meaningful contrasts in their input. This hypothesis is tested across five typologically-diverse languages that vary by phone inventory size and structure as well as word composition. Consonant FL was calculated over more than 390,000 words of child-directed speech. Results demonstrate that FL correlates positively with earlier consonant emergence in all languages. Models fit to bootstrapped corpus data include both FL and frequency as predictors, but suggest that frequency may be the stronger of the two. A need to complicate assumptions on the relationship between environmental effects and phonological development is discussed.

1. Introduction

There are many predictors of when a child will first produce a consonant: articulatory complexity, input frequency, language phonotactics, and even word structure. Though all likely contribute to emergence, the relative influence of each factor may vary by language. For example, with an articulation towards the front, and at times back, of the mouth, the lateral approximant /l/ is articulatorily complex. It is, accordingly, late to emerge in English (Lin & Demuth, 2015). However, laterals emerge fairly early in Quiché Mayan, as soon as 1;7, a fact which Pye, Ingram, & List (1987) attribute to high frequency in the ambient language.

Beyond frequency, a more precise predictor of consonant emergence may be functional load (FL). FL is the entropy loss a system undergoes due to phone convergence (Hockett, 1955). In its basic form, FL measures how many minimal word pairs a sound distinguishes (e.g. pat~bat). Interest in FL as an explanatory device for sound mergers and inventories has recently resurfaced (Surendran & Levow, 2006; Wedel, Jackson, & Kaplan, 2013). It has also been applied, to a limited extent, as a predictor of child consonant development (Pye et al., 1987; Stokes & Surendran, 2005; Van Severen et al., 2013). And its potential usefulness as a metric of phone emergence is often suggested, though not implemented (So & Dodd, 1995) or conflated with frequency (Amayreh, 2003). Frequency is a useful predictor of consonant emergence and mastery (Edwards & Beckman, 2008; Edwards, Beckman, & Munson, 2015) and it is a natural correlate with FL. However unlike frequency, FL encompasses semantic contrast and lexicon structure. This may be critical for development since lexicon-derived phonetic categories, over

---

* For feedback and assistance, thank you to the fall 2017 members of Cognitive Science 190: Computational Models of Cognition (UC Berkeley), especially Michael Levy. Additional thanks to Chris Hench and Ronald Sprouse.

those purely inferred from distributions in the input, have resulted in more robust category acquisition for models of infant learners (Feldman, Griffiths, & Morgan, 2009).

The relationship between FL and phonological development is intuitive. As children selectively focus upon contrasts in their ambient language, they acquire phones that frequently generate semantic contrasts first (Dietrich, Swingley, & Werker, 2007). Still, it is not clear if the contrastive importance of a phone always reinforces consonant emergence. Data from several unrelated languages show conflicting evidence. Like the relationship between the emergence and articulatory complexity of /l/, the explanatory power of FL may differ cross-linguistically. For example, Stokes & Surendran (2005) attributed the low predictive power of FL in Cantonese consonant development to word structure: with six tones, Cantonese has many segmental homonyms and this could lower the FL of individual phones. But Van Severen et al. (2013) found that FL was a better predictor of consonant emergence in Dutch than frequency alone. Thus how individual languages mitigate the roles of FL and frequency for development remains unclear. Finally, there are multiple approaches to FL calculation – its role on phonological emergence has not been uniformly evaluated across multiple languages. A standard analysis of FL calculation can address the universality of FL for consonant emergence.

Understanding the varied parameters of speech development has clear clinical implications for normally- and atypically-developing children. But is the contribution of FL and frequency the same for all children? Or does it depend on the language of exposure? Furthermore, though the negative correlation of FL and consonant emergence is intuitive, the reason why a child prioritizes highly contrastive phones in phonological learning is less straightforward. In fact, this distributional learning mechanism may differ by language. Here this is addressed by computing FL and phone frequency over the child-directed speech (CDS) of five typologically-diverse languages that vary by phone inventory size and structure and word composition. This cross-linguistic comparison evaluates the potentially language-specific role of FL and frequency in child consonant emergence.

2. Methods

2.1. Corpora preprocessing

FL and frequency were calculated over naturalistic, monolingual corpora of American English (Bernstein-Ratner [Bernstein-Ratner, 1987] & Brent-Ratner corpus [Brent & Cartwright, 1996]), Japanese (MiiPro corpus [Miyata, 2012]), Shenzhen Mandarin (Tong corpus [Deng & Yip, 2017]), Peninsular Spanish (Aguirre corpus [Aguirre, 2000]), and Turkish (Aksu [Slobin, 1982] & Altinkamis corpus [Türkay, 2005]) available in CHILDES (MacWhinney, 2000). These languages were selected because they were either 1) already phonologically transcribed or 2) relatively orthographically transparent which permits algorithmic grapheme-to-phoneme conversion. Only CDS directed towards the child from 1;0-3;0 and from the target child's mother, father, grandparents, and adult interlocutors was included. Though sibling input undoubtedly impacts development, sibling utterances were excluded since age and presence was corpus-dependent. To further increase corpus generalizability, the following were also removed: all proper nouns, with the exception of familial terms (e.g. "Mama"), child- and family-specific forms, second language items, and investigator speech. This resulted in the following token counts/corpus: English (N=32,993), Japanese (N=235,705), Mandarin (N=72,908), Spanish (N=44,440), and Turkish (N=10,977). Discrepancies in corpus size are counteracted in a bootstrap procedure before model fitting in *Results*.

The Mandarin (Pinyin transcription), Spanish, and Turkish corpora underwent a grapheme-to-phoneme conversion utilizing the Montreal Forced Aligner (McAuliffe et al., 2017). Forms without a representation in the corresponding dictionary were discarded. These unknown words made up 0.39%, 0.86%, and 10.71% of the Spanish, Mandarin, and Turkish corpora, respectively. Tone was removed from the Mandarin corpus. The Brent-Ratner corpus for English was already transcribed phonologically and the MiiPro Japanese corpus was transcribed in Hepburn Romanization which is orthographically transparent (Miyata p.c.).

2.2 Calculation

FL is defined as the system entropy loss resulting from phoneme convergence/loss. It can be formalized as follows:

$$FL_U(a) = \frac{CL_U - C(L_U^{-a})}{CL_U}$$

where *a* is the linguistic unit, *C* is entropy, and L_U is the linguistic system. Both FL and frequency were calculated independently for consonants over corpora word types (English [N=1,321], Japanese [N=10,414], Mandarin ([N=2,200], Spanish [N=2,305], and Turkish [N=2,216]). This study focuses on system entropy at the word level but it should be noted that phoneme-level entropy is another useful calculation for FL (Surendran & Niyogi, 2006).

FL was measured over the entire consonant inventory of each language (Table 1) except 1) Japanese geminate stops[†], 2) Japanese /ʃ/ since there was not any developmental data, 3) Spanish /θ/ and /s/ since the developmental data reports non-*ceceo* dialects, and 4) nasals which are ubiquitous from early babbling and likely emerge too early to be lexically meaningful. Anterior voiced stops are also early to emerge. However these data are included to contrast with the emergence of voiced velars and voiceless/aspirated stops. English /ʒ/, Spanish /x/, and Turkish /ɣ/ (“soft g”) and /ʒ/ had a 0 FL and are not included in further analyses.

There is disagreement concerning best practices for FL calculation. Stokes & Surendran (2005) justify the choice to calculate FL only in word-initial position since “children pay attention to the onsets of words (581).” However, this is not universal. For example, in early word production, French children actually tend to omit word-initial segments, likely due to exclusive word-final stress in French (Vihman, 2013). Consequently, here FL is calculated over all segments. Elsewhere, FL calculations are limited to the lemma (Wedel et al., 2013). But since Turkish, a highly agglutinating language, is included in this analysis, all inflected and derived forms in the remaining languages are also. Finally, the frequency of a phone is defined as the number of its occurrences in the corpus divided by the number of total phones in the corpus.

2.3 Developmental data

Age of consonant emergence (AoE) was determined from previous peer-reviewed works of developmental phonology for each language. Studies have employed distinct metrics to qualify a consonant as “emerged” in a child’s phonological repertoire: if the sound was present in the child’s inventory at least two times in a given speech sample, for example (Dinnsen et al., 1990) or, in larger studies, if 90% of the children produced the sound one time (Prather, Hedrick, & Kern, 1975; So & Dodd, 1995). Data collection methodologies – naturalistic, elicited, etc. – also

---

[†] Emergence was limited to diary data, not acoustic. See Kunnari, Nakai, & Vihman (2001) for acoustic data on Japanese geminate mastery.

differ by study. Table 2 lists studies referenced and the metrics employed. When an age range was specified (e.g. 18-22 months), the mean month was taken as AoE.

LANGUAGE	STOPS	AFFRICATES	FRICATIVES	LIQUIDS/GLIDES
English	p, t, k, b, d, g	tʃ, dʒ	f, v, θ, ð, s, z, ʃ	l, ɹ, w, y
Japanese	p, t, k	ts, tʃ	s, z, h,	r, w, y
Mandarin	p, t, k, p^{h}, t^{h}, k^{h}	ts, ts^{h}, tʂ, tʂ^{h}, tɕ, tɕ^{h}	s, f, ʂ, ɕ, x	ɻ, l
Spanish	p, t, k, b, d, g	tʃ	f	l, ɾ, r, ʝ
Turkish	p, t, k, b, d, g	tʃ, dʒ	s, f, v, ʃ, ɣ, h	l, r, j

Table 1: Consonants measured

Given the discrepancy between studies, the metric for consonant emergence is not standard. For example, consonants appear to emerge much later in English but this is due to the more stringent emergence criterion used in Prather et al. (1975). As a result, AoE is not directly comparable between languages – only its relation to FL and frequency.

3. Results

Figure 1 maps the relationship of normalized FL to age of emergence (AoE). To compare between languages, FL was normalized by the sum of all FL calculations within each language: $FL(a)$ / $\sum Fx_i$. The negative correlation of FL and AoE for each language varied from mild to high: Japanese (Pearson r = -.28), English (r = -.46), Mandarin (r = -.34), Turkish (r = -.48), and Spanish (r = -.91) meaning that as children get older, the consonants that emerge in their speech tend to have a lower FL. Correlations did not include outliers (English /s, z/, Mandarin /ts/, Spanish /ɾ/, Japanese /k, t/). These are marked in gray in the figures. FL correlated more strongly with AoE than frequency for four out of five languages: English (r =-.42), Japanese (r = -.07), Mandarin (r = -.45), Spanish (r = -.83), Turkish (r = -.15). This suggests that FL may play a strong role in consonant emergence for some languages, though frequency could still be predictive.

To confirm the generalization from correlational statistics, a bootstrapping with replacement procedure was employed over each of the CDS type-frequency language corpora in one hundred 750-word samples.[‡] FL was then normalized over the sum of phone FL measurements within each sample.

AoE was binned into developmental periods of 3 months (e.g. 0;11-1;1, 1;2-1;5) since it does not evoke a continuous period of development, rather discrete, chronologically-ordered developmental stages. A stepwise cumulative link mixed effects model, similar in implementation and interpretation to other hierarchical (e.g. mixed effects) models but specified for ordinal response variables, was fit with the `clmm` function in the R package `ordinal` (Christensen, 2015). Multivariate linear and even logistic models cannot perform as well as cumulative link models. Linear models predict values outside of the realistic range of the response variable. Here this means that the model would predict a relationship before and after the age that children begin producing consonants. Forwards stepwise model fitting was evaluated

---

[‡] In Japanese, smaller word samples resulted in a FL=0 for almost all segments/sample. This is likely due to heterogeneous phonotactics and distinct lexical strata in Japanese so a larger sample size is required to gauge FL (Itô & Mester, 1999).

through log-likelihood tests and AIC comparison. Generalizations about the relationship between emergence and FL likely do not extend to segments with large FLs so the outliers /s, z/ (English), /ts/ (Mandarin), /ɾ/ (Spanish), and /k, t/ (Japanese) were not included in model fitting.

LANGUAGE	REFERENCE	METRIC FOR EMERGENCE
English	Prather et. al (1975)	75% of children (*N*=147) used consonant in initial and final position
Japanese	Nakanishi (1982); Ota (2015)	mean age of first appearance across (*N*=10) children
Mandarin	Hua & Dodd (2000)[§]	90% of children in age group (~*N*=20) produced sound once
Spanish	Cataño, Barlow, & Moyna (2009) (metanalysis)	occurred at least twice in given speech sample from *N*=16 children
Turkish	Topbaş (1997)	produced in all seven possible word positions by *N*=22 children

Table 2: Metric for consonant emergence

When discussing child phonology, an obvious concern about the role of motor limitations arises. No model of phonological emergence is complete without a metric of articulatory complexity but it is surprisingly complex to quantify. The scale of the model parameter **articulatory_complexity** used here is adopted from Stokes & Surendran (2005) (Table 3). To ensure convergence, articulatory_complexity stands as a proxy for individual phones.

Best model fit included both **FL** (β= -1.41630, ***) and **frequency** (β = -75.25579, ***) as well as **articulatory_complexity** and random slopes for **language*articulatory_complexity**. Articulatory complexity alone does not explain when children will begin to produce consonants; both frequency and FL are important parameters for models of consonant emergence, at least for phones that do not have exceptionally large contrastive power. Negative coefficients signify that as children age, phone FL and frequency decrease. Children learn higher FL and higher frequency phones first, even when controlling for articulatory demands. Finally, FL and frequency metrics were transformed to z-scores to test variable importance: FL: β= -0.08332, Frequency: β= -1.17255. With these coefficients, we can conclude that frequency is the stronger predictor of emergence, though both are vital components to the final model.

4. Discussion

Phone frequency influences consonant development in children's speech. The more children hear a sound, the faster they can focus attention to imitating its articulation and ensuing acoustic signal. Yet this intuitive relationship has limitations. Edwards et al. (2015) cite the example of

---

[§] Hua & Dodd (2000) examine phonological development in Putonghua Mandarin (standard Beijing) while the Mandarin CDS corpus documents development in Shenzhen Mandarin, a southern variety.

English /ð/ – a phone with exceptionally high token frequency due to words like ‘the’ and ‘that’, but low type frequency – that emerges relatively late in development. So while frequency is

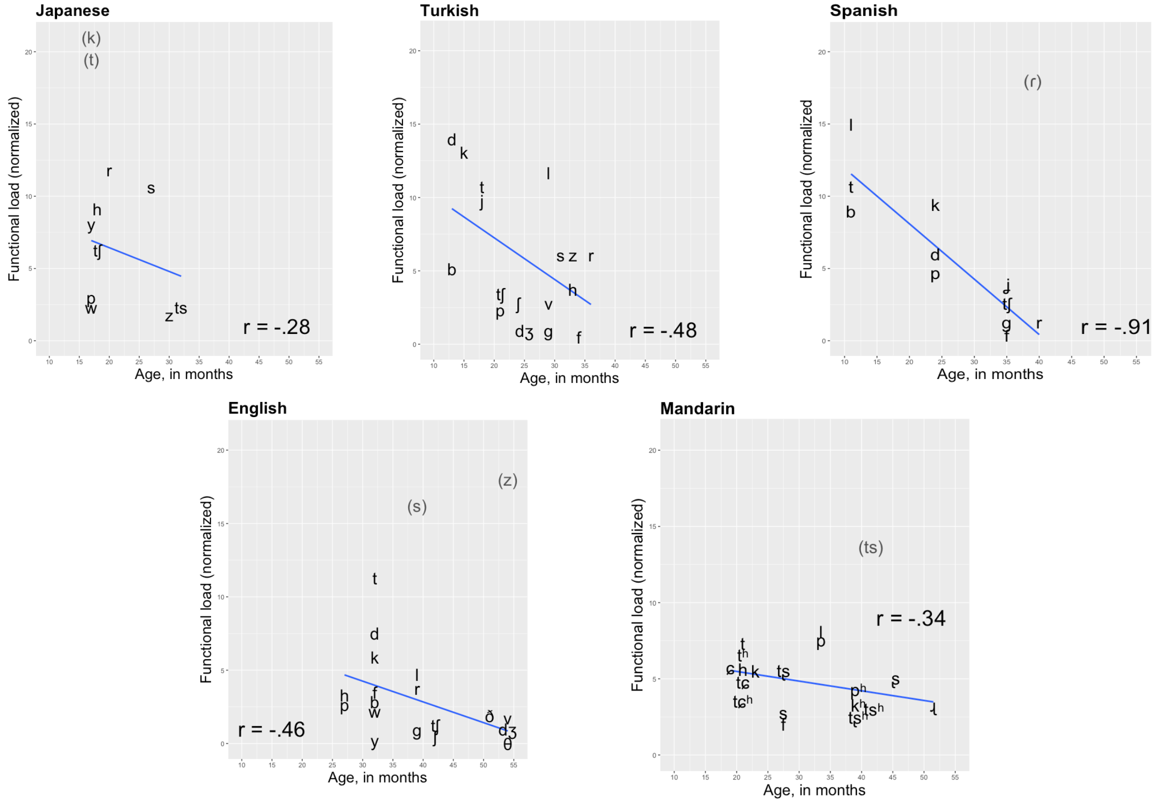

Figure 1: Functional load by age of emergence across languages

LEVEL	DESCRIPTION	PHONES
1	-Rapid/ballistic movement -Slow/progressive movement	p, w, h
2	-Some lingual control for frication -Velar place of articulation -Laryngeal mastery	k, t, b, g, j, f, d, pʰ, tʰ, kʰ
3	-Tongue tip and dorsum manipulation	r, ɹ, ɻ, ɾ, l
4	-Complete lingual control for fricatives -Transition from ballistic to frication	ʃ, ð, θ, v, ʝ, x, tʃ, dʒ, ts, tsʰ, ʂ, tʂ, tʂʰ, ɕ, tɕ, tɕʰ

Table 3: Articulatory complexity metric
(Adapted with modification from Stokes & Surendran [2005])

correlated with development, alone, it cannot complete a picture of phonological developmental. Likewise, child phonology is rife with examples of phones that emerge late due to motor limitations and immature physiology (McGowan, Nittrouer, & Manning, 2004). But cross-linguistically, the same segment can emerge in child speech at different developmental stages (Edwards & Beckman, 2008). So articulatory demands also do not fully predict when sounds emerge in child speech.

The model here incorporates both of these factors and tests an additional parameter that contributes to consonant emergence across all languages tested: functional load. Even when controlling for type frequency and articulatory complexity, children manipulate the semantic information in the input to inform the timeline of early phone and word productions. This supports previous findings about the primacy of the lexicon for phonological development (Feldman et al., 2009).

There is of course a natural cyclicity to any argument about ambient language effects and language acquisition. The argument follows a chicken-or-the-egg logic: do children acquire segments because they are more frequent in the input or more frequently contrast words? Or are those sounds more frequent because they are more "naturally" acquired or easier to pronounce? Both explanations are valid and this model suggests that, cross-linguistically, universals and language-specific parameters govern phonological development.

The conclusion regarding the constant relationship between FL and emergence across all the tested languages warrants discussion. It suggests that we may need to alter our current understanding of environmental effects on consonant emergence. The few studies to test a relationship between consonant emergence and FL have not found evidence that it uniformly predicts consonant emergence. Here both type frequency and FL contribute to emergence, but Stokes & Surendran (2005) concluded that FL predicts consonant emergence in English to the complete exclusion of frequency. Likewise, Van Severen et al. (2013) state that FL impacts emergence more than frequency alone for children acquiring Dutch. Yet as previously mentioned, Stokes & Surendran (2005) also concluded that FL has an inconsequential effect upon emergence in Cantonese, potentially due to the contrastive power of tone in the language. Of course there is no evidence of this for Mandarin, the tonal language tested here.

The differences between studies could be due to methodological choices. For example, Stokes & Surendran (2005) calculated the FL of Cantonese over a corpus of child speech, not child-directed. (Van Severen et al. [2013] demonstrated that FL calculated over child-directed speech is a better correlate with emergence than adult-directed speech). Second, this study estimated that highly frequent/contrastive phones may not follow a predictable relationship between emergence and FL. So unlike previous studies, phones with a high FL were excluded from model fitting. Calculations were also performed over type, not token, corpora. Finally, the developmental data was rightly treated ordinally. Future analyses into the relationship between emergence and ambient effects should manipulate each of these factors in turn.

In conclusion, the intuition that ambient frequency predicts phonological development was confirmed. Likewise the ability to contrast words, calculated by a phone's FL, also predicts early speech production. This relationship is constant across the five languages studied, English, Japanese, Mandarin, Spanish, and Turkish and maintained even after controlling for the articulatory demands of phones. These results reaffirm the dual contributions of environment and physiology on early consonant production.

5. References

Aguirre, C. (2000). La adquisición de las categorías gramaticales en español. Ediciones de la Universidad Autónoma de Madrid.

Amayreh, M. M. (2003). Completion of the consonant inventory of Arabic. *Journal of Speech, Language, and Hearing Research*, *46*(3), 517-529.

Bernstein-Ratner, N. (1987). The phonology of parent-child speech. In K. Nelson and A. van Kleeck, editors, *Children's Language*, volume 6. Erlbaum, Hillsdale, NJ.

Brent, M.R. and Cartwright, T.A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* 61: 93-125.

Cataño, L., Barlow, J. A., & Moyna, M. I. (2009). A retrospective study of phonetic inventory complexity in acquisition of Spanish: Implications for phonological universals. *Clinical Linguistics & Phonetics*, *23*(6), 446-472.

Christensen, R. H. B. (2015). Ordinal - Regression Models for Ordinal Data. R package.

Deng, X. & Yip, V. (2017). A multimedia corpus of child Mandarin: the Tong corpus. *Journal of Chinese Linguistics.*

Dietrich, C., Swingley, D., & Werker, J. F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *PNAS*, *104*(41), 16027-16031.

Dinnsen, D. A., Chin, S. B., Elbert, M., & Powell, T. W. (1990). Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics. *Journal of Speech, Language, and Hearing Research*, *33*(1), 28-37.

Edwards, J., & Beckman, M. E. (2008). Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. *Language Learning and Development*, *4*(2), 122-156.

Edwards, J., Beckman, M. E., & Munson, B. (2015). Frequency effects in phonological acquisition. *Journal of Child Language*, *42*(2), 306.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). Learning phonetic categories by learning a lexicon. *Proceedings of the 31st Annual Conference of the Cognitive Science Society.*

Itô, J. and Mester, A. (1999). The structure of the phonological lexicon, in N. Tsujimura (ed.), *The handbook of Japanese linguistics*. Blackwell.

Jarosz, G. (2010). Implicational markedness and frequency in constraint-based computational models of phonological learning. *Journal of Child Language*, 37(3), 565-606.

Kunnari, S., Nakai, S., & Vihman, M. M. (2001). Cross-linguistic evidence for the acquisition of geminates. *Psychology of Language and Communication*, *5*(2), 13-24.

Lin, S., & Demuth, K. (2015). Children's acquisition of English onset and coda /l/ : Articulatory evidence. *Journal of Speech, Language, and Hearing Research*, *58*(1), 13-27.

Hockett, C. F. (1955). *A manual of phonology* (No. 11). Waverly Press.

Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language*, *27*(1), 3-42.

McAuliffe, M. Socolof, M., Mihuc, A., Wagner, M., Sonderegger, M. (2017). Montreal Forced Aligner [Computer program]. Version 1.0.0.

McGowan, R. S., Nittrouer, S., & Manning, C. J. (2004). Development of [ɹ] in young, Midwestern, American children. *JASA*, *115*(2), 871-884.

Miyata, S. (2012). Japanese CHILDES: The 2012 CHILDES manual for Japanese.

Nakanishi, Y. (1982). K_on hattatsu. In R. Uchisugawa and Y. Yasuko (Eds.) *K_za gengo sh_gai chiry_ky_iku 4: K_onsh_gai*. Tokyo: Fukumura Shuppari.

Ota, M. (2015). L1 phonology: phonological development. In H. Kubozono (Ed.), *Handbook of Japanese Language and Linguistics, vol. 2*. Berlin: Mouton de Gruyter.

Prather, E. M., Hedrick, D. L., & Kern, C. A. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders*, *40*(2), 179-191.

Pye, C., Ingram, D. & List, H. (1987). A Comparison of Initial Consonant Acquisition in English and Quiché, in Keith Nelson and Anne van Kleeck (Eds.), *Children's Language*, *Vol. 6*. Erlbaum.

Slobin, D. (1982). Universal and particular in the acquisition of language. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art*. New York: Cambridge University Press.

So, L. K., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language*, *22*(3), 473-495.

Stokes, S. F., & Surendran, D. (2005). Articulatory complexity, ambient frequency, and functional load as predictors of consonant development in children. *Journal of Speech, Language, and Hearing Research*, *48*(3), 577-591.

Surendran, D., & Niyogi, P. (2006). Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. *Amsterdam studies in the theory and history of linguistic science series* 4, 279, 43.

Topbas, S. (1997). Phonological acquisition of Turkish children: implications for phonological disorders. *European Journal of Disorders of Communication*, *32*(4), 377-396.

Türkay, A. F. (2005). *Children's Early Lexicon in terms of Noun-Verb Dominance*. Unpublished PhD Dissertation. Cukurova Univ., Adana, Turkey.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, *128*(2), 179-186.

Van Severen, L., Gillis, J. J., Molemans, I., Van Den Berg, R., De Maeyer, S., & Gillis, S. (2013). The relation between order of acquisition, segmental frequency and function. *Journal of Child Language*, *40*(4), 703-740.

Vihman, M. M. (2013). *Phonological development: The first two years*. John Wiley & Sons.