



HAL
open science

Markedness, Frequency: Can We Predict the Order of Acquisition of Consonants?

Naomi Yamaguchi

► **To cite this version:**

Naomi Yamaguchi. Markedness, Frequency: Can We Predict the Order of Acquisition of Consonants?. LingO 2007, the second Oxford Postgraduate Conference in Linguistics, Sep 2007, Oxford, United Kingdom. pp.236-243. halshs-00472649

HAL Id: halshs-00472649

<https://shs.hal.science/halshs-00472649>

Submitted on 12 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Markedness, Frequency: Can We Predict the Order of Acquisition of Consonants?

NAOMI YAMAGUCHI

Laboratoire de Phonétique et Phonologie

(UMR 7018 – CNRS / Université Paris III Sorbonne Nouvelle)

naomi.yamaguchi@univ-paris3.fr

ABSTRACT.

Several recent studies have discussed the role of frequency in the acquisition of phonemic contrasts. The aim of this paper is to test two competing hypotheses on data composed of productions from 26 French-speaking children, aged from 3;07 to 4;05. The first hypothesis is that phoneme frequency determines the order of acquisition of consonants. The second one is that feature frequency determines the order of acquisition of consonants. Then, we will show that feature frequency is a better predictor than phoneme frequency in our data. Finally, we will propose that feature frequency can define the traditional notion of markedness.

1 Introduction

In studies of language acquisition, the most frequently asked question is “how does a child acquire a specific language?” Several answers to this question have been proposed, and one of them is based on frequency. It has been shown that infants are sensitive to statistics in early perception, (Anderson, Morgan, and White 2003), as well as in production (Zamuner, Gerken, and Hammond 2005; Stites, Demuth, and Kirk 2004).

In this paper we will compare two different hypotheses, stated as follows:

- (1) Phoneme frequency in adult speech predicts the order of acquisition of consonant phonemes for French children.
- (2) Feature frequency in adult speech predicts the order of acquisition of consonant phonemes for French children.

2 Methodology

Child data

The children’s productions come from the Chevrie-Muller and Plaza (2001) corpus.

Markedness, Frequency: Can We Predict the Order of Acquisition of Consonants?

The data used here are taken from 26 monolingual French children (13 boys and 13 girls), aged between 3;07 and 4;05. None of the children were undergoing speech therapy.

The data collection was carried out according to an experimental protocol, consisting in a questionnaire of words, with all the French consonant phonemes in all positions, except /ɲ/ only found in coda position. Each child was in a classroom alone with the reviewer (a speech therapist). The reviewer showed images of objects well-known by children (cat, dog, cake) or pointed to parts of her body, and asked the child what the object was. If the child didn't answer, the reviewer tried to elicit the name of the object, using semantic prompts, by giving a definition or citing examples (without phonemic cues). Then, if the child didn't know the answer, the reviewer said the name of the object and asked the child to repeat it. For the present study we only based our analyses on the spontaneous productions, and not the repetitions.

Four different transcribers transcribed the data in broad phonetic transcription: the first reviewer (the speech therapist), two linguistics students, and the author. The transcribers were trained in phonetic transcription. To encode the data, the transcription used was the one with at least 75% of agreement (3 transcribers out of 4).

Table 1 shows the consonant phonemes considered to be totally acquired by children, meaning that they are correctly realized in each occurrence. Thus /p, t, n/ are the consonants totally acquired by all children; /k, m, l, j/ were totally acquired by all but one; and so forth.

Number of children	Consonantal phonemes
26	p, t, n
25	k, m, l, j
24	b, d, g, f, s, ʁ
23	v, z
18	ʃ
16	ʒ
15	ɲ

Table 1: **Totally** acquired consonants

It will be assumed here that the rank order shown in this table corresponds to the order of acquisition for this sample of children.

Frequency lists

In order to test the hypothesis in (1), we compared consonant frequencies found in two spoken language corpora to the consonant rankings shown in Table 1. The first corpus, shown in table 2, from Adda-Decker (2006), shows consonant frequencies in adult data. The data is taken from telephone conversations between adults.

The second frequency list (table 3) used in this study is taken from Le Calvez

Rank	Phoneme	Proportion (in %)
1	ʁ	5,86
2	l	5,76
3	s	5,41
4	t	5,06
5	m	4,73
6	k	4,05
7	p	3,62
8	d	3,45
9	v	2,41
10	n	2,23
11	ʒ	1,89
12	b	1,45
13	f	1,42
14	j	1,17
15	z	1,05
16	ʃ	0,53
17	g	0,38
18	ɲ	0,05

Table 2: Consonant frequency in adult data (Adda-Decker 2006)

Rank	Phoneme	Proportion (in %)
1	s	6,05
2	l	6,02
3	ʁ	5,98
4	t	5,91
5	k	4,68
6	p	3,99
7	m	3,21
8	d	3,10
9	n	2,32
10	v	2,27
11	b	1,43
12	f	1,38
13	j	1,35
14	ʒ	1,08
15	ʃ	0,92
16	g	0,64
17	z	0,42
18	ɲ	0,01

Table 3: Consonant frequency in CDS data (Le Calvez 2004)

(2004). Le Calvez made her list on French corpora from the CHILDES database, on child directed speech (henceforth CDS) produced by adults. The orthographic transcriptions were phonetically encoded using VoCoLex (Dufour, Peereman, Pallier, and Radeau 2002).

These lists show us that the most frequent consonants in French are /t, s, l, ʁ/, in adult data as well as in CDS data. Moreover, the two lists don't differ from each other in a significant way (tested with the statistical test of Chi square; $\chi^2 = 16,84$ with $p = 0,465$).

3 Phoneme frequency

If the first hypothesis (in (1)) is true, then we can verify this influence by carrying out two kinds of analyses. First, a quantitative approach can be used, by means of statistical tests; second, a qualitative approach can be taken, by comparing predictions with the observed data.

Statistical testing

If phoneme frequency is related to consonant acquisition, then it will be related quantitatively. We ran a Spearman correlation test, comparing phoneme frequencies in each of our corpora to the ranked consonant data in our child inventories. The correlation is described in figures 1 and 2.

The correlation coefficient between the adult frequency list and the consonant rankings in child inventories is: $\rho = 0,538$ with $p < 0,05$ and the correlation coefficient between the CDS frequency list and the consonant rankings in child inventories is: $\rho = 0,626$ with $p < 0,01$.

Markedness, Frequency: Can We Predict the Order of Acquisition of Consonants?

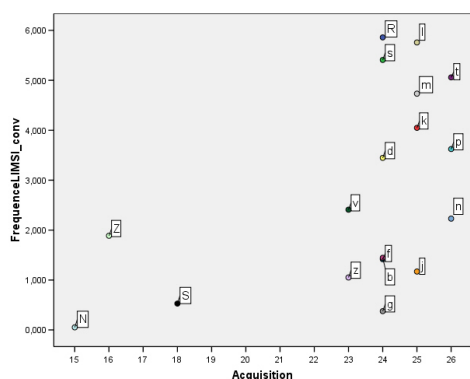


Figure 1: Correlation using the adult frequency list

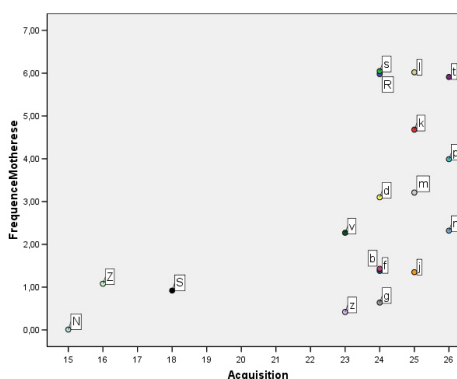


Figure 2: Correlation using the CDS frequency list

These results show that in each case, while there is some correlation (0 = no correlation and 1 = full correlation), the coefficients are quite weak. Moreover, the fact that there is a statistical correlation doesn't mean that one variable (phoneme frequency) is the cause for the second variable (consonant acquisition in child inventories). This correlation test just shows that there is a relation between phoneme frequency and consonant acquisition, which may be caused by another factor.

Comparing predictions with the observed data

The quantitative tests show a relation between phoneme frequency and consonants acquisition in child inventories, but they haven't proved the influence of phoneme frequency on the order of acquisition of consonants. To examine this question further, we carried out a qualitative analysis based on the predictions of the phoneme frequency model.

The prediction based on phoneme frequency can be stated as follows:

- (3) *the most frequent member of each minimally distinct pair of consonants will be the first acquired.*

The CDS corpus will be used for testing this prediction. All the predictions following from the frequencies shown in table 3 for French consonants are listed in table 4.

In table 4, all the rows in bold face show incorrect predictions, in which the first phoneme to be acquired by the children is the less frequent member of the pair. For instance, let's consider the pair /f/ – /v/. According to the phoneme frequency list in table 3, /v/, ranked at 10th position, is more frequent than /f/, ranked at 12th position. So, according to the prediction in (3), /v/ should be acquired before /f/, and should appear in more inventories than /f/. Looking at the child inventories, we see that the opposite is true: /f/ appears in more inventories than /v/.

Feature	Pairs of phonemes	Prediction by frequency	First acquired phoneme in child inventories
[voiced]	p – b	p	p
	t – d	t	t
	k – g	k	k
	f – v	v	f
	s – z	s	s
	j – ʒ	ʒ	j
[PLACE]	p – t	t	same number of inventories
	t – k	t	t
	p – k	k	p
	b – d	d	d
	b – g	b	same number of inventories
	b – ɣ	b	same number of inventories
	f – s	s	same number of inventories
	v – z	v	same number of inventories
	m – n	m	n
	l – ʁ	l	l
[continuant]	p – f	p	p
	t – s	s	t
	b – v	v	b
	d – z	d	d
	l – j	l	same number of inventories
[posterior]	s – ʃ	s	s
	z – ʒ	ʒ	z
	n – ɲ	n	n
[lateral]	n – l	l	n

Table 4: Predictions made by phoneme frequency

There are 8 cases like this one, where phoneme frequency is in conflict with the acquisition inventories. These results show that phoneme frequency makes incorrect predictions about the order of consonantal phonemes in acquisition.

4 Feature frequency

We have shown that phoneme frequency does not make correct predictions about the order of consonantal phonemes. Given that phonemes are organized according to features, we may turn to our second hypothesis, according to which feature frequency in adult speech predicts the order of acquisition of consonant phonemes for French children¹. We test the predictions of this hypothesis below. Data is drawn again from the CDS corpus (table 3).

Phonemes are organized in a system according to contrasts, which are represented in terms of features. For instance, /p/ contrasts with /b/ in terms of the feature [voiced], and the same feature makes a further contrast between /t/ and /d/, and between /k/ and /g/.

Feature frequency calculation

A feature value frequency is calculated by adding the frequencies of all consonants in which this feature value is distinctive.

¹I would like to thank John Coleman for this relevant proposal

Markedness, Frequency: Can We Predict the Order of Acquisition of Consonants?

Feature	Relevant class	Values	Phonemes	%	Most frequent
[voice]	[-son]	[-voice]	p, t, k, f, s, ʃ	22,93	[-voice]
		[+voice]	b, d, g, v, z, ʒ	8,94	
[CORONAL] ^a	[+cons]		t, d, s, z, ʃ, ʒ, n, l, j, ɲ	27,18	[COR]
[DORSAL]			k, g ɣ	11,30	
[LABIAL]			p, b, f, v, m	10,01	
[continuant]	[-son]	[-cont]	p, t, k, b, d, g	19,75	[-cont]
		[+cont]	f, s, ʃ, v, z, ʒ	12,12	
[posterior]	[COR]	[-post]	s, z, n	8,79	[-post]
		[+post]	ʃ, ʒ, ɲ	2,01	
[sonorant]	[+cons]	[-son]	p, t, k, b, d, g, f, s, ʃ, v, z, ʒ	31,87	[-son]
		[+son]	m, n, ɲ, l, ɸ ^b	18,89	

^a[CORONAL], [LABIAL], [DORSAL] are one-valued features, forming the set of place features in which [CORONAL] is the most frequent member.

^b/ɸ/ is considered as a sonorant given diachronic and phonotactic reasons.

Table 5: Feature value frequencies

For instance, to calculate the frequency of [+voiced], we add the frequencies of /b, d, g, v, z, ʒ/, excluding sonorants for which [voiced] is not distinctive.

We obtain the frequencies of all feature values shown in table 5.

Predictions

The prediction made by hypothesis (2) (based on feature frequency) is that the phoneme bearing the most frequent feature value will be the first acquired. We now compare the predictions of this hypothesis with those of the preceding one. The table 6 is given in the same format as table 4, but adds the predictions made by feature frequency in the fourth column.

There are 7 cases of disagreement between the predictions made by phoneme frequency and those made by feature frequency as shown by the bold face rows. In all these cases, the predictions made by feature frequency are in agreement with the child inventories.

There is one case (the shaded row), where both phoneme frequency and feature frequency make incorrect predictions about child inventories. The prevalence of the /p/ in the face of the greater frequency of /k/ may be explained by other factors: there may be a visual reinforcement of the labial, for instance.

5 General discussion and conclusion

Phoneme frequency makes incorrect predictions about consonant acquisition; whereas feature frequency makes much better predictions. First of all, these results confirm the role of features in acquisition, rather than phonemes.

Features play a role in acquisition, and specifically some feature values tend to be preferred. Looking closer at the most frequent feature values, we obtain [-voi-

Feature	Pairs of phonemes	Prediction by phoneme freq.	Prediction by feature freq.	First acquired phoneme in child inventories
[voice]	p – b	p	p	p
	t – d	t	t	t
	k – g	k	k	k
	f – v	v	f	f
	s – z	s	s	s
	ʃ – ʒ	ʒ	ʃ	ʃ
[PLACE]	p – t	t	t	same nb of inventories
	t – k	t	t	t
	p – k	k	k	p
	b – d	d	d	same nb of inventories
	b – g	b	g	same nb of inventories
	b – d	b	d	same nb of inventories
	f – s	s	s	same nb of inventories
	v – z	v	z	same nb of inventories
	m – n	m	n	n
	l – ʀ	l	l	l
[continuant]	p – f	p	p	p
	t – s	s	t	t
	b – v	v	b	b
	d – z	d	d	d
	l – j	l	l	same nb of inventories
[posterior]	s – ʃ	s	s	s
	z – ʒ	ʒ	z	z
	n – ɲ	n	n	n
[lateral]	n – l	l	n	n

Table 6: Predictions made by phoneme frequency and made by feature frequency

ced], [CORONAL], [-continuant], [-posterior]. The remarkable fact is that these feature values correspond exactly to the feature values considered as *unmarked* in the phonological literature. (See e.g. Brandao de Carvalho 1997; Clements 2007). Markedness and frequency, more precisely probability of occurrence in the world’s languages, are related in a typological way (Greenberg 1966; Martinet 1965). According to Clements (2007, ip):

- (4) “Within any class of sounds in which a feature is potentially distinctive, the marked value of this feature is [the one which is] absent in some languages and the unmarked value of this feature is [the one which is always] present.”

By this definition, the feature values which are unmarked in the world’s languages, i.e. those that are present in all languages, include [-voiced], [CORONAL], [-continuant], [-posterior]. It is striking to notice that these are exactly the same feature values that are the most frequent in French, at least within the CDS corpus.

This result confirms a strong link between frequency and markedness. We suggest that feature frequency could be the basis for a more a precise definition of the controversial notion of markedness.

References

- Adda-Decker, M. (2006). De la reconnaissance automatique de la parole à l'analyse linguistique de corpus oraux. In *Actes des XXVI^{es} journées d'études sur la parole*, pp. 389–400.
- Anderson, J., J. Morgan, and K. White (2003). A statistical basis for speech sound discrimination. *Language and Speech* 46, 155 – 182.
- Brandao de Carvalho, J. (1997). Primitives et naturalité. *Langages* 31, 14–34.
- Chevrie-Muller, C. and M. Plaza (2001). *Nouvelles Epreuves pour l'Examen du Langage (N-EEL)*. Paris: Les Editions du Centre de Psychologie Appliquée.
- Clements, G. (2007). L'évitement de la marque : vers un nouveau modèle des inventaires phonologiques. In E. Delais-Roussarie and L. Labrune (Eds.), *Des sons et des sens : données et modèles en phonologie et en morphologie*. Paris & London: Hermès.
- Clements, G. (i.p.). The role of features in speech sound inventories. In E. Raimy and C. Cairns (Eds.), *Contemporary Views on Architecture and Representations in Phonological Theory*. Cambridge, MA: MIT Press.
- Dufour, S., R. Peereman, C. Pallier, and M. Radeau (2002). Vocolex : Une base de données lexicales sur les similarités phonologiques entre les mots français. *L'année psychologique* 102, 725 – 746.
- Greenberg, J. (1966). *Language Universals* (2ème édition 2002 ed.). Berlin: Mouton de Gruyter.
- Le Calvez, R. (2004). Modélisation de l'acquisition des catégories phonémiques. Master's thesis, EHESS, Paris.
- Martinet, A. (1965). *La linguistique synchronique : études et recherches*. Paris: Presses Universitaires de France.
- Stites, J., K. Demuth, and C. Kirk (2004). Markedness vs frequency effects in coda acquisition. In A. Brugos, L. Micciulla, and C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development*, pp. 565–576.
- Zamuner, T. S., L. A. Gerken, and M. Hammond (2005). The acquisition of phonology based on input: a closer look at the relation of cross-linguistic and child language data. *Lingua* 10, 1403–1426.