

## The Emergent Syllable

**John J. Ohala**

*Department of Linguistics*

*University of California, Berkeley*

*Berkeley, California, USA*

It is a well-known principle in science that theories must be falsifiable or, as some have put it “unless a theory can be mortally threatened, it cannot live.” It is in this spirit that I present the following arguments against the Frame/Content theory about the primacy of jaw oscillations and of the syllable in human speech. I therefore assume the role similar to “the loyal opposition” in the British Parliament. Similarly to other theories, the Frame/Content theory should not be passively accepted without a thorough weighing of alternatives.

### **The Historical Context of How Evolutionary Theories are Evaluated**

A problem for all the sciences of evolution and development—whether phylogenetic or ontogenetic—is how to raise the standard of evidence in support of a theory on origins of an organism or behavior to a level where reasonable people can give it allegiance over competing theories. Darwin’s theory of evolution via the natural selection of variants is an example. Experiments seemed not to be possible and quantitative data are difficult to obtain, unlike the

case with physics and chemistry. (Today, however, quantification of genes largely overcomes the latter difficulty.) The answer to this issue is strengthening the theory by citing multiple case studies and interlocking point-by-point similarities. A single case study showing a posited pattern is not particularly convincing but 12, 25, 50 such cases showing the same pattern gives increasing credibility to the reality of the pattern. (Boë (this volume) shows how, in ideal cases, the necessary number of similarities can be estimated probabilistically. Darwin's *The origin of species* (1859) as well as his *The expression of the emotions in man and animals* (1872) are such large, fat, volumes because they are, for the most part, filled with case studies. The method of multiple case studies is one of the cornerstones of many sciences where measurements are difficult, e.g., evolution of behaviors, epidemiology, and historical linguistics.

It is instructive to examine the development of this method in historical (comparative) linguistics. Before this scientific breakthrough there were centuries of speculation on family relationships between languages, on the history of particular words, on cognate words in different languages (see, e.g., ten Kate (1723), de Broses (1765), Burnet (Lord Monody) (1773-1792). But there was very little rigor in establishing the similarities between words. Voltaire is reputed to have remarked, sarcastically, that in etymology "[similarity in] consonants count little, and vowels nothing." What the classical grammarians (Gyarmathi, Bopp, Rask, Grimm, among others) did, was to note point-by-point similarities and relationships among large numbers of words.

When comparing the candidate cognate words Latin *pater* and Gothic *fadar*, e.g., they pointed out similarities (e.g., the labiality and voicelessness of the initial consonants) as well as differences (e.g., stop in Latin but fricative in Gothic). The same pattern of similarities and differences appear in many other cognate pairs as exemplified in Table 8.1, where English cognates are given instead of Gothic. From this analysis, these early grammarians were able to extract the generalization that a voiceless stop in Latin corresponds to a homorganic fricative in Gothic (with certain positional exceptions). The number

**Table 8.1.**

A small sample of cognate words in Latin and English.

Latin	Germanic (English)
pisces	Fish
ped	Foot
tenuis	Thin
tres	Three
centum	Hundred
cornua	Horn
cannabis	Hemp

of regular interlocking similarities exceeded what any reasonable person would say could have originated by chance.

It is very difficult to achieve this level of confidence in any of the variety of interesting and colorful theories proposed on the origin of language and speech especially, I would maintain, on the relevance of concepts like syllabicity.

### Chewing as the Seed of Speech

Regarding the origin of speech there have been various proposals that speech precursors are to be found in:

- chewing (Froeschels, 1951; MacNeilage, 1998; Weiss, 1950)
- sound imitation (de Broses, 1765, and many others)
- iconic gestures of the tongue, lips, etc. made audible (Fonagy, 2001; Paget, 1923)
- physiologically common or necessary sounds (Müller, 1861)
- emotional sounds (cries) that were decoupled from their original stimulus, perhaps for purposes of deception (Müller, 1861)
- and dozens of others

Only one of these places any emphasis on action of the jaw originally associated with chewing.

My point is not to say that the articulations original to chewing were not the precursors to speech but rather to say that there is no way to evaluate this story vis-à-vis the others. All of them have their merits and demerits. As I pointed out elsewhere (Ohala, 1998), there are some points of dissimilarity between chewing and speaking, e.g., in chewing there is a significant lateral movement of the jaw that is missing from speech.

### The Importance of Acoustic Modulations in Speech

But my primary misgivings about putting so much emphasis on jaw movements is not to say that they are unimportant but rather that they are subordinate to—or dependent on—an even more important principle of communication using the vocal-auditory channel (or any other channel, for that matter). This principle is that all signals must show syntagmatic contrast, i.e., modulation (Ohala, 1995).

In speech these modulations may be and are in any of several parameters: amplitude, fundamental frequency (F0), spectrum (location of spectral peaks, bandwidths of these peaks, spectral tilt), source characteristics (voicelessness, voicing, creaky voice, breathy voice, etc.), and duration (insofar as certain articulations have an expected or default duration, any deviation from the default constitutes a potential modulation or contrast that can be useful in signaling). Certainly movements of the jaw create modulations in amplitude and spectrum and, in some cases, periodicity, depending on the degree of closure of the jaw. But jaw movement by itself neglects F0 and neglects the fact that other articulatory gestures, especially those of the larynx, velum, the tongue, and the lips (without the help of the jaw) can create such modulations. Would anyone seriously want to argue that laryngeal or source modulations played a secondary role in early hominid vocal communication? For that matter, would anyone seriously want to claim that source (laryngeal) modulations play a secondary role in early vocalizations of human infants? Crying, without rhythmic oscillations of any articulator, is the earliest vocalization exhibited by infants. As every

parent recognizes, babies early on learn to de-couple crying from actual physical distress and use it to garner special benefits from caregivers. Is this a precursor of vocal symbolization? I make no claims on this point but the idea has as much superficial plausibility as the claim that the gestures found in chewing are the precursors of speech. Some authors have claimed that the first contrastive, i.e., truly linguistic, vocal signals in children involve use of intonation (e.g., Smith, N. V., 1973). When humans developed the cognitive capacity to recognize the signaling possibilities of vocal communication, i.e., when the vocal signals could symbolize and refer to something, is it not plausible that the usefulness of the source modulations would immediately present themselves, especially as these were certainly already familiar to them, being part of the vocal repertory of all known species using vocal-auditory communication (virtually all mammals, birds, some amphibians, some fish)? For the earliest vocal communication then, the syllable seems too advanced.

### The Syllable: An Ill-defined Entity

In any case, I believe the notion of the syllable is too ill defined to propose it as a cornerstone of speech.

There have been numerous phonetically discredited definitions or “findings” regarding the nature of the syllable:

- Separate breath pulses that is, pulsatile contractions of the respiratory muscles (Stetson, 1928; but see Ladefoged, 1967; Ohala, 1990).
- Local maxima in vocal tract aperture. This proposal fails to account for the supposed monosyllabicity of words like *spa* or the bi-syllabicity of words in tone languages like [n̄má] or the bi- vs tri-syllabicity of contrasting words like English *lightning* vs. *lightening*.
- Local maxima in sound amplitude. Some of the above words constitute exceptions to this proposal, too.

The notion of “sonority” is often invoked in discussions of syllables: speech sounds are claimed to be arranged at the beginning of syllables in the

order of increasing sonority and (perhaps) at the end of syllables in the order of decreasing sonority. The sonority scale (from least to most) is posited to be:

stops – fricatives – nasals – “liquids” – glides – vowels

Thus /ta tra mla smja/ obey this principle; /fta mta lma/ do not. But sonority is an empirically empty concept that can no more be determined for speech sounds than can their temperature (Ohala & Kawasaki-Fukumori, 1997). There is a viable alternative—I return to the idea that what matters is syntagmatic contrast.

Two fundamental changes—vis á vis the supposed sonority hierarchy—are proposed to account for syntagmatic phonological constraints, as outlined in Table 8.2. From this latter view there is nothing anomalous about English *spa*, Mandarin [sz] or [s], words like [spsk] in Nez Perce. It also explains why sequences like the following are avoided in many languages even though they are easily articulated and obey the supposed sonority constraint: [ji wu bw-]. The reason: they make insufficient syntagmatic contrast.

In this new view the syllable is not a basic organizing principle. It is, rather, epiphenomenal; some language-specific organization of a *collection of parametric modulations* (of acoustic variables). Changes in the direction of modulation of one or more (especially more) of these parameters may be taken as “peaks” or “valleys” in the sequence of sounds. There is no plan behind them except, perhaps, for the weight of such patterns in other words in the languages lexicon. In this view the ambiguity in number of syllables in words like English *Brian* vs. *brine*, *towel* vs. *owl*, is to be expected: some of the modulations in the sound sequences are rather subtle.

I hasten to add, however, that what I’ve said does not imply that the syllable has no reality. It could be an emergent entity, a grouping that speakers impose on the stream of speech just as, for example, experienced typists impose on frequent letter sequences. If that is true it does not detract from the fundamental organizing principle in vocal-auditory signal systems of syntagmatic contrast which supercedes and may encompass whatever is meant by syllabicity.

**Table 8.2.**

Comparison of the constraints on sound sequencing. In the case of the proposed constraints under the “RIGHT” heading, the concept of the syllable is unnecessary.

WRONG	RIGHT
Trying to identify a single parameter, sonority, whose empirical correlates are elusive (or non-existent)	Identifying several parameters whose empirical correlates are well known, e.g., the acoustic parameters amplitude, periodicity, spectrum, F0, duration
Claiming that segments arrange themselves according to their inherent value of this parameter	Claiming that segments are arranged according to the relative differences in these parameter. The more these parameters differ between adjacent segments, the better and thus the more commonly they are found in the languages of the world.

## Conclusion

### My main points are these:

1. The Frame/Content theory that the rhythmic action of the jaw as seen in chewing is the basic frame for articulate speech is interesting but there is no compelling reason to give it any credence.
2. The claim that syllables are somehow the basic unit of speech neglects the difficulty of defining the syllable and neglects the possibility that the more fundamental principle governing speech structure is syntagmatic contrast, including especially source modulations.

## References

- Bloomfield, L. (1933). *Language*. New York: Holt.  
 Boë, L. J., Bessiere, P., Ladjili, N., & Audibert, N. (2008). Simple combinatorial considerations challenge Ruhlen’s mother tongue theory. In B. L. Davis & K.

- Zajdo (Eds.), *The syllable in speech production* (pp. 63–93). London: Taylor & Francis.
- de Brosses, Ch. (1765). *Traité de la formation mécanique des langues, et de principes physiques de l'étymologie*. 2 vols. Paris: Chez Saillant, Vincent, Desaint.
- [Burnet, J. = Lord Monboddo]. (1773–1792). *Origin and progress of language*. (6 vols.) Edinburgh (various printers).
- Darwin, C. (1859). *The origin of species*. London: John Murray.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- Fonagy, I. (2001). *Languages within language: An evolutive approach*. Amsterdam: John Benjamins.
- Froeschels, E. (1952). Chewing method as therapy. *AMA Archives of Otolaryngology*, 56, 427–434.
- ten Kate, L. (1723). *Aenleiding tot de Kennisse van het verhevene Deel der nederduitsche Sprake*. 2 vols. Amsterdam: Rudolph en Gerard Wetstein.
- Ladefoged, P. (1967). *Three areas of experimental phonetics*. Oxford: Oxford University Press.
- MacNeilage, P. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21, 499–511.
- Müller, M. (1861). Lectures on the science of language, delivered at the Royal Institution of Great Britain in April, May, and June, 1861. London: Longman, Green, Longman, and Roberts.
- Ohala, J. J. (1990). Respiratory activity in speech. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 23–53). Dordrecht: Kluwer.
- Ohala, J. J. (1995). Speech perception is hearing sounds, not tongues. *Journal of the Acoustical Society of America*, 99, 1718–25.
- Ohala, J. J. (1998). Content first, frame later. *Behavioral and Brain Sciences*, 21, 525–527.
- Ohala, J. J., & Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In S. Eliasson & E. H. Jahr (Eds.), *Language and its ecology: Essays in memory of Einar Haugen* (pp. 343–365). Berlin: Mouton de Gruyter.
- Paget, R. (1930). *Human speech*. London: Kegan, Paul, Trench, Trubner & Co.
- Smith, N. V. (1973). *The acquisition of phonology*. Cambridge: Cambridge University Press.
- Stetson, R. H. (1928). *Motor phonetics: A study of speech movements in action*. [Archives néerlandaises de phonétique expérimentale, Tome 3].
- Weiss, D. A. (1950). Chewing and the origin of speech. In: H. H. Beebe & D. A. Weiss (Eds.), *The chewing approach in speech and voice therapy* Sasel: Karger.
- Young, T. (1819). Remarks on the probabilities of errors in physical observations, and on the density of the earth, considered, especially, with regard to the reduction of experiments on the pendulum. *Philosophical Transactions of the Royal Society of London*, 109, 70–95.

## Co-occurrence Patterns in the Babbling of Children with a Cochlear Implant \*

Karen Schauwers<sup>1,2</sup>

Paul J. Govaerts<sup>1,2</sup>

and

Steven Gillis<sup>1</sup>

<sup>1</sup> Department of Linguistics

University of Antwerp

Antwerp-Wilrijk, Belgium

<sup>2</sup> The Eargroup

Antwerp-Deurne, Belgium

Prelexical babbling represents an important achievement in children's vocal development. Although the characteristics of babbling (i.e., onset, segmental content) have been studied intensively in the last twenty years or so, it still remains unclear to what extent this prelexical development is autonomous or driven by auditory input and feedback.

---

\* The research reported in this paper was supported by a grant from the Science Foundation – Flanders, contract G.0216.05.

# THE SYLLABLE IN SPEECH PRODUCTION

Barbara L. Davis • Krisztina Zajdó

## Contents

Foreword by Björn Lindblom	ix
Preface by Barbara L. Davis and Krisztina Zajdó	xxi
List of Contributors	xxxix

### Overview

1. The Frame/Content Theory <i>Peter F. MacNeilage</i>	1
<b>I: Evolutionary Perspectives on Speech Development</b>	
2. The Origins of Syllabification in Human Infancy and in Human Evolution <i>D. Kimbrough Oller and Ulrike Griebel</i>	29
3. Simple Combinatorial Considerations Challenge Ruhlen's Mother Tongue Theory <i>Louis-Jean Boë, Pierre Bessière, Nadia Adjili, and Nicolas Audibert</i>	63
4. The Frame/Content Theory and the Emergence of Consonants <i>Didier Demolin</i>	93
5. Lipsmacking and Babbling: Syllables, Sociality, and Survival <i>John L. Locke</i>	111

Lawrence Erlbaum Associates  
Taylor & Francis Group  
270 Madison Avenue  
New York, NY 10016

Lawrence Erlbaum Associates  
Taylor & Francis Group  
2 Park Square  
Milton Park, Abingdon  
Oxon OX14 4RN

© 2008 by Taylor & Francis Group, LLC  
Lawrence Erlbaum Associates is an imprint of Taylor & Francis Group, an Informa business

Printed in the United States of America on acid-free paper  
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-0-8058-5480-0 (Softcover) 978-0-8058-5479-4 (Hardcover)

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at  
<http://www.taylorandfrancis.com>

and the LEA and Routledge Web site at  
<http://www.routledge.com>