

Language Faculty and Beyond *Internal and External Variation in Linguistics*

Language Faculty and Beyond (LFAB) focuses on research that contributes to a deeper understanding of the properties of languages as a result of the Language Faculty and its interface with other domains of the mind/brain. While the series will pay particular attention to the traditional tension between descriptive and explanatory adequacy, the series will also address issues such as the level of linguistic design, through new lines of inquiry often referred to as 'physiological linguistics' or 'biolinguistics'. LFAB aims to publish studies from the point of view of internal and external factors which bear on the nature of micro- and macro-variation as, for example, understood in the minimalist approach to language.

For an overview of all books published in this series, please see
<http://benjamins.com/catalog/lfab>

Editors

Kleanthes K. Grohmann
University of Cyprus

Pierre Pica
CNRS, Paris

Advisory Board

Paola Benincà
University of Padova, Italy

Cedric Boeckx
ICREA/University of Barcelona, Spain

Guglielmo Cinque
University of Venice, Italy

Noam Chomsky
Massachusetts Institute of Technology,
Cambridge, USA

Stephen Crain
Macquarie University, Sydney, Australia

Marcel den Dikken
CUNY Graduate Center, New York, USA

Naama Friedmann
Tel Aviv University, Israel

Anders Holmberg
University of Newcastle, UK

Lyle Jenkins
Biolinguistics Institute, Cambridge, USA

Richard K. Larson
Stony Brook University, USA

Andrew Ira Nevins
University College London, UK

Alain Rouveret
University of Paris VII, France

Esther Torrego
University of Massachusetts, Boston, USA

Anna Papafragou
University of Delaware, Newark, USA

Akira Watanabe
University of Tokyo, Japan

Volume 12

The Form of Structure, the Structure of Form. Essays in honor of Jean Lowenstamm
Edited by Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi
and Nicola Lampitelli

The Form of Structure, the Structure of Form

Essays in honor of Jean Lowenstamm

Edited by

Sabrina Bendjaballah
CNRS

Noam Faust
Hebrew University of Jerusalem

Mohamed Lahrouchi
CNRS

Nicola Lampitelli
Université de Tours

John Benjamins Publishing Company
Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

The form of structure, the structure of form : essays in honor of Jean Lowenstamm /

Edited by Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi and Nicola Lampitelli.

p. cm. (Language Faculty and Beyond, ISSN 1877-6531 ; v. 12)

Articles in English, Spanish, and French.

Includes bibliographical references and index.

1. Afroasiatic languages--Grammar. 2. Grammar, Comparative and general--Phonology, Comparative. 3. Grammar, comparative, and general--Morphology. 4. Afroasiatic languages--Phonetics. I. Faust, Noam, editor. II. Lahrouchi, Mohamed, editor. III. Lampitelli, Nicola, editor. IV. Lowenstamm, Jean, honouree.

PJ993.F67 2014

492'.045--dc23

2014027028

ISBN 978 90 272 0829 3 (Hb ; alk. paper)

ISBN 978 90 272 6948 5 (Eb)

© 2014 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Introduction	1
PART 1. Vowels	
Lowering harmony in Bantu: An RcvP account <i>Harry van der Hulst</i>	13
On vowel harmony and vowel reduction: Some observations on canonical shapes of disyllabic nouns in Yukuben, Mòoré and German <i>John R. Rennison</i>	37
Apophony and chiming words in Malay <i>Gilles Boyé</i>	57
Understanding what has happened with the ablaut: A chapter from linguistics <i>Gabriel Bergounioux</i>	67
PART 2. Syllables	
On the licensing of glides <i>Mohand Guerssel</i>	85
Coda constraints on tone <i>Larry M. Hyman</i>	103
C/V interactions in strict CV <i>Joaquim Brandão de Carvalho</i>	123
What does the Moroccan Malhun meter compute, and how? <i>Mohamed Elmedlaoui</i>	139

PART 3. Templates

Regularities in irregular Chaha verbs <i>Degif Petros Banksira</i>	163
Overlapping morphologies in Arabic hypocoristics <i>Jean-François Prunet and Ali Idrissi</i>	177
Staying away from the weak left edge: A strengthening strategy <i>Outi Bat-El</i>	193
The Modern Hebrew template tQuLa in light of Jean Lowenstamm's work <i>Noam Faust</i>	209
Templates and representations in phonology: From Semitic to child language <i>Sophie Wauquier</i>	219
On templates <i>Alain Kihm</i>	235
PART 4. The Phonology-Syntax interface	
The Ins and Outs of phonology <i>Jonathan Kaye</i>	255
Phase cycles, ϕ -cycles, and phonological (In)activity <i>David Embick</i>	271
Sepp vs Paradigms <i>Sabrina Bendjaballah and Martin Haiden</i>	287
On Plurals, noun phrase and num(ber) in Moroccan Arabic and Djibouti Somali <i>Mohamed Lahrouchi and Nicola Lampitelli</i>	303
The initial CV: Herald of a non-diacritic interface theory <i>Tobias Scheer</i>	315

PART 5. Selected Issues in Afro-Asiatic (Morpho-)Syntax and Semantics

Causatives, anticausatives and lexicalization <i>Jamal Ouhalla</i>	333
A note on labeling, Berber states and VSO order <i>Ur Shlonsky</i>	349
The interpretation of Construct-State morphology <i>Edit Doron</i>	361
Index	375

2. Somali examples are written according to the official spelling.
3. The suffix *-e* becomes *-a-* when it appears in an intervocalic position. The same holds for the suffix *-ó*, cf. *naagó* 'women' vs. *naagáyáal* 'groups of women'.
4. It should be mentioned that such double formatives occur in the plural and not in the singular. As far as we know, none of the languages discussed above has double singular forms. This may relate to the standard issue of the input to derivation, well-known in morphological theory: is the plural form derived from the singular or from an abstract root? Due to space limitations, we leave the issue aside. The reader is referred to Prunet (2006) for a detailed review of the literature.
5. The suffix *-at* happens to be the feminine plural marker, mainly in adjectives and participials: for example, *zwin* 'beautiful (M SG)'/*zwinin* (M PL)/*zwinat* (F PL). However, this does not hold for nouns since *-at* does not necessarily change the gender of the noun it combines with.
6. The diacritics " ' " indicates the accent pitch. Cf. Hyman (1981) for further details. Somali has four syntactic cases: ABS(solutive), nominative, genitive and vocative. The nouns given in (5) display the default case, i.e. ABS (cf. Saeed 1993).
7. The plural of nouns in (5b) is characterized by the gemination of the last consonant of the stem.
8. The endings *-é* and *-ó* are suffixes, as noted by Puglielli and Siyaad (1984, 59–63) and later by Lecarme (2002). Cf. *bár* 'to teach' → *baré* 'teacher'.
9. Lampitelli (2013), building on Godon (1998), proposes that the unique plural suffix of Somali nouns is *-o*. This suffix has three allomorphs: *-o* (*naagó*, *albaabbó*, *ilkó*), *-a-* (*miisás*) and zero (*madáx*). The accentual pitch is the marker of the syntactic case.
10. In DJ, an M noun displays at least two different ways to pluralize: (1) the suffixation of *-yáal* and (2) the Standard Somali pluralization pattern to which it belongs (cf. 5).
11. According to Puglielli and Siyaad (1984, 92–93), the plural of plural suffix is *-yow* (*-yáal* is not mentioned). The suffix *-yow* is always masculine.

The initial CV

Herald of a non-diacritic interface theory

Tobias Scheer

Université de Nice – Sophia Antipolis, CNRS 7320

Across all structuralist and generative incarnations, interface theories are based on unwarranted diacritics: #, ω etc. Lowenstamm's initial CV has paved the way for a non-diacritic interface theory: the carrier of morpho-syntactic information in phonology is syllabic space. The article shows that the initial CV is able to kill three birds with one stone (strength of initial consonants, initial cluster restrictions, syncope of first vowels). It is then argued for the idea that carriers of morpho-syntactic information in phonology reduce to syllabic space (there are no other carriers). Finally, it is shown that boundary information does not belong to the same lexical entry as morphemic information: the initial CV for example is not present in the lexical entry of the word it precedes.

1. Introduction

Lowenstamm (1999) argues that the phonological identity of the beginning of the word is syllabic space, an empty CV unit. The purpose of the pages below is to show three things:¹

1. The specific implementation of the initial CV in the environment of CVCV theory provides an answer to a puzzle where three birds need to be killed with one stone. This puzzle has not even been identified by other theories.
2. The beginning of the word is but one particular case of what syllabic space marks: empty CV units also spell out all other kinds of morpho-syntactic information. Carriers of morpho-syntactic information in phonology reduce to syllabic space.
3. The phonological string is pieced together by two kinds of objects: those coming from vocabulary (or lexical) insertion (morphemic information), and those which represent on line-created morpho-syntactic information

(boundary information). Just like all other boundary information, the initial CV is not a piece of the lexical entry of morphemes (or words). It sits in a lexical entry of its own.

The take-home message is then that Lowenstamm's initial CV has paved the way for a more general and, as we will see, non-diacritic interface theory that offers conceptual and empirical advantages. Without exception and across structuralist as well as various generative incarnations, existing interface theories are all based on unwarranted diacritics.

2. Diacritic sleepers vs. phonologically meaningful objects

Since the 19th century (at least), phonologists know that phonology cannot be done without specific reference to the beginning of the word. This reference was always made to a diacritic, whose identity was determined by the basic phonological units of the time: boundary "phonemes" in structuralism when phonology was a string of phonemes, # which was held to be a [-segment] segment in SPE when the basic phonological units were feature matrices, and finally autosegmental domains such as the prosodic word ω or the prosodic phrase φ since the early 80s when all areas of phonology were autosegmentalized.

Hence the interface vocabulary used in all theories was only ever diacritic. The initial CV argued for by Lowenstamm (1999) is not: syllabic constituents, an onset and a nucleus, are not arbitrarily chosen and interchangeable symbols whose function reduces to the representation of morpho-syntactic information. If a # or an ω is replaced by a banana and phonological processes are then said to occur in the vicinity of a banana, rather than of a # or an ω , the workings of the interface as we know it will remain unchanged. By contrast, a CV unit cannot be replaced by a banana for the trivial reason that it has a phonological identity which is independent from its eventual function of carrying morpho-syntactic information: an onset is an onset, not a banana; a nucleus is a nucleus, not a banana.

A number of arguments can be made against diacritic carriers of morpho-syntactic information (see Pak 2008, 60ff; Samuels 2009, 284ff; Scheer 2008; 2012, §93). Below the most obvious of them is recalled: diacritics are intrinsically unable to make predictions. They are empty shells since they do not carry any information except the morpho-syntactic load that they are designed for.² By contrast, non-diacritics such as CV units do make precise predictions. Beyond the analytic virtue that the initial CV originally was designed for, Lowenstamm (1999) has thus introduced a new interface vocabulary: non-diacritic items.

In order to see the difference, consider that phonology does not react on the simple presence of a hash mark or a prosodic word – such items can only bear on phonology if the analyst has implemented some instruction in phonological computation (a rule or a constraint) that is sensitive to them. That is, hash marks, omegas and the like are colourless and passive "sleepers". They merely sit in phonological representations without producing any effect by themselves.

For example, will a hash mark or a prosodic word favour or disfavour consonant clusters in their vicinity? There is no answer to this question because they can trigger (or inhibit) any phonological process and its reverse. By contrast, if phonologically meaningful vocabulary carries morpho-syntactic information, phonological computation will react on its bare presence. This is what I call the Direct Effect: the simple *presence* of phonologically meaningful objects makes predictions. Consider the following example.

- (1) equally probable rules?
 a. $V \rightarrow \emptyset / \#C_CV$
 b. $\emptyset \rightarrow V / \#C_CV$

Both rules under (1) are equally probable and equally natural from the point of view of a theory that uses diacritics: no property of the theory favours or disfavors the epenthesis into an initial cluster, or the deletion of a vowel in this context. Every phonologist knows, though, that (1b) is an attested phonological process, while (1a) is not on record. That is, there is no "masochistic" language that would delete vowels in initial clusters (and only in this context).³

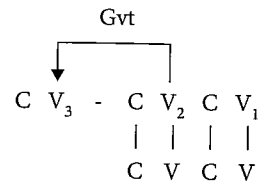
Therefore theories that cannot discriminate between (1a) and (1b) have a problem, and the reason why they are in trouble is that the critical information, i.e. word-initiality, is conveyed by a diacritic hash mark. The result is the same in case the prosodic word or some other prosodic constituent carries this information: anything and its reverse may happen at the left edge of a prosodic constituent. Prosodic constituents are sleepers without any intrinsic property that produce an effect only when some computational instruction (such as *#... or *[ω ...] calls on them.

The two rules at hand are discriminated as soon as the extra-phonological information "beginning of the word" comes as a real phonological object whose mere *presence* impacts phonology and which does not need to be explicitly mentioned in rules (or constraints) in order to produce an effect.

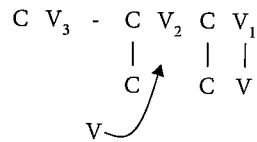
Table (2) depicts the two situations in the environment of CVCV (Lowenstamm 1996, Scheer 2004), with the beginning of the word incarnating as an empty CV unit.

(2) deletion vs. insertion of the first vowel in a word in CVCV

a. deletion: ill-formed



b. insertion: structure rescued



Being a phonologically meaningful object, the leftmost CV unit under (2 a,b) that represents the beginning of the word is parsable by phonological computation. Under (2b) where the first nucleus of the root is empty, its presence creates a sequence of two empty nuclei. Since V_1 can only govern V_2 , V_3 will remain ungoverned, which means that the structure is ill-formed. Therefore an epenthesis into V_2 will rescue the word.

On the other hand, the structure under (2a) is well-formed: V_2 governs V_3 , and no empty nucleus remains orphan. The deletion of the content of V_2 , however, creates a sequence of two empty nuclei and therefore makes the structure ill-formed: this is the “masochistic” option.

It is thus predicted that the deletion rule (2a) is impossible, while the epenthesis rule (2b) is regular – exactly what we find across languages. Therefore there is a clear difference between non-predictive diacritics which allow anything and its reverse to happen in their vicinity, and truly phonological objects that have a predictable effect on the well-formedness of phonological structure.

3. The initial CV parameterized: Languages may or may not have it

3.1 Stable cross-linguistic effects disqualify diacritics

Since the bare presence of diacritics does not have any effect and makes no prediction, representing boundary information by diacritics is claiming that anything and its reverse can be triggered by them: nothing predestines a # to provoke, say, the presence or the absence of consonant clusters in its vicinity, and even less so the presence or absence of a specific kind of cluster (say, TR).⁴

The diacritic-based prediction is incorrect: the phonological phenomena that are observed in this context are anything but arbitrary (Scheer 2004, §§83, 402; 2009a; 2012, §246; Ségéral and Scheer 2008b). Languages may or may not grant a specific status to the word-initial site. If they do not, word-initial consonants behave just as if they were word-internal. If on the other hand something peculiar happens at left edges of words, it is always the same phenomena that are observed. Table (3) below lists three phenomena that are recurrently found to occur at the left edge of words.

(3) stable effects of the beginning of the word across languages

a. restrictions on initial clusters

in some languages initial CCs are restricted to #TR. In others they have the same distribution as internal CCs. But there is no language where they are restricted to #RT.

b. strength of initial consonants

in some languages word-initial consonants are especially strong. In others, they do not have any peculiar behaviour regarding strength. But there is no language where they are especially weak.

c. deletion of the first vowel of the word

in some languages the first vowel of words is unable to alternate with zero. In others it does not show any peculiar behaviour with respect to other vowels. But there is no language where non-initial vowels are unable to alternate with zero, while initial vowels are.

Hence in case the word-initial position is outstanding (which it may or may not be), it allows only for a characteristic subset of logically possible events. Therefore diacritics are out of business. What we need is a truly phonological exponent of the boundary information “beginning of the word” whose substance predicts just the effects observed, and no others.

Below it is shown that the three cross-linguistically stable effects of the beginning of the word are explained (and their reverse excluded) if the initial CV, couched in the lateral network of CVCV, is assumed.

3.2 Three for the price of one

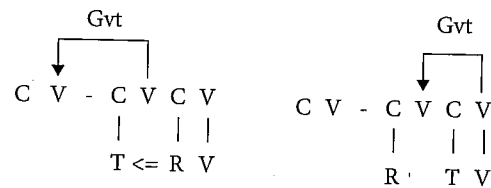
3.2.1 TR-only vs. anything-goes languages.

Let us start by looking at restrictions on word-initial clusters. It was shown in Scheer (2004, §§102, 402; see also Scheer 1999; 2007) that the presence of the initial CV restricts word-initial clusters to #TR. This is because the empty nucleus which separates TR clusters, but not the one in RT (or TT, RR) clusters, is silenced by Infrasegmental Government (<= under (4a)). Being ungoverned, \emptyset in T \emptyset R is thus a good lateral actor and able to govern.⁵

By contrast, the two members of an RT cluster cannot contract Infrasegmental Government (Scheer 2004, §14); therefore \emptyset in $R\emptyset T$ needs to be governed by the following nucleus as under (4b) and is thus unable to dispense government. As a consequence, #RT clusters are ill-formed in presence of an empty nucleus to their left since two empty nuclei in a row are illegal. On the other hand, #TR clusters are well-formed because the empty nucleus enclosed is able to govern the empty nucleus that comes with the initial CV.

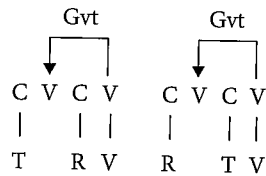
(4) restrictions on initial clusters in CVCV languages that possess the initial CV

- a. #TR well-formed b. #RT ill-formed: two empty nuclei in a row



languages that lack the initial CV

- c. #TR and #RT well-formed



The pattern discussed thus far is the typical Indo-European situation found in English, Italian, French and the like where word-initial clusters are restricted to obstruent-liquid. This is what may be called TR-only, and we have seen that this restriction on word-initial clusters is enforced by the presence of the initial CV.

There are other languages, however, where non-TR clusters occur word-initially: Moroccan Arabic, Greek or Polish for example instantiate the anything-goes pattern. Given the analysis of TR-only languages, it may be calculated what it takes for an initial #RT cluster to be well-formed: the initial CV must be absent. This is shown under (4c): since there is only one empty nucleus to be silenced, the first vowel of the word can always do the job no matter what the sonority slope of the cluster. Hence any sequence of two consonants can exist word-initially: anything goes in languages that lack the initial CV.⁶

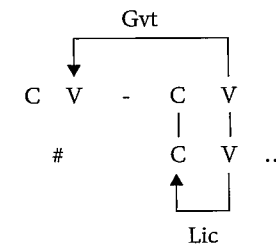
3.2.2. Alternating first vowels and the strength of initial consonants.

In some languages, first vowels of words cannot alternate with zero (while vowels elsewhere in the word can). Relevant evidence is discussed in Ségéral and Scheer (2008b) and Scheer (2004, §90). The reason why first vowels cannot alternate with zero is the same as before: the relevant configuration is shown under (4b), only that the governed nucleus accommodates a vowel-zero alternation. Zero surfaces under government, which creates a situation where the presence of the initial CV produces a sequence of two empty nuclei, which is ill-formed. Therefore first vowels of words resist vowel-zero alternations in languages that mark the left edge of words with the initial CV.

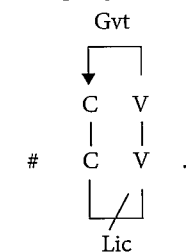
Turning to lenition and fortition, a central observation is that in some languages word-initial consonants are strong. In the environment of the Coda Mirror (Ségéral and Scheer 2001; 2008b; Scheer 2004, §110), the strength of word-initial consonants is a consequence of the presence of an empty nucleus to their left: the government of the following vowel is absorbed by the initial CV, which means that the consonant itself is licensed (i.e. backed up) but ungoverned (i.e. unspoiled), that is, experiences maximally comfortable conditions (under (5a)).

(5) strength of word-initial consonants

- a. languages with the initial CV



- b. languages without the initial CV



(5b) shows the situation of word-initial consonants in languages where no CV unit is distributed at the left edge of words. As may be seen, the first vowel of the word is now free to govern its own onset, which therefore remains unlicensed (in terms of the Coda Mirror v2: see Scheer and Ziková 2010; Scheer 2012, §223). In such a

system, initial consonants thus experience intervocalic conditions. That is, consonants are not only predicted to be weak in languages without the initial CV – the prediction is more precise: word-initial consonants in these languages stand in intervocalic (rather than in coda) position. This is indeed the case in Greek (see Seigneur-Froli 2006; Scheer 2012, §333).

4. Predictions made by the parameterisation of the initial CV

Given the preceding, the parameterisation of the initial CV predicts that any language which displays one of the three properties of the righthand or the lefthand column under (6) below must also instantiate the two other properties of the same column.

(6) consequences of the parametric presence/absence of the initial CV	
initial CV present	initial CV absent
a. word-initial consonants are strong	word-initial consonants are non-strong
b. word-initial clusters are restricted to #TR	there are no restrictions on word-initial clusters: #TR, #RT, #TT and #RR may freely occur
c. first vowel of words may not alternate with zero	first vowel of words may alternate with zero

These predictions are anything but trivial: they chain together three empirical situations that otherwise seem to be unrelated. Also, they are empirically explicit and may be easily falsified.⁷

The examination of languages that I am familiar with produces encouraging results, even though most of the time there are only valid diagnostics for two of the three criteria. Regarding left-column languages for example, initial consonants are always strong in typical #TR-only languages such as Romance and Germanic (see Ségéral and Scheer 2008a). In the right column, in many (but not all) Slavic languages such as Polish, Czech or Russian, initial #RT clusters occur (the Slavic situation is discussed at greater length in Scheer 2007; 2012, §320). In the same languages, the first vowel of words may alternate with zero (e.g. Czech *lev* – *lv-a* ‘lion Nsg, Gsg’). Also, Kijak (2005) shows that in Mazovian Polish (a North-Eastern dialect), the only case in Slavic that I am aware of where a diagnostic for the strength of word-initial consonants is available, these appear to be weak.⁸

By contrast, in a left-column Slavic language such as Belarusian that imposes the #TR-only restriction on initial clusters, the typical Slavic vowel-zero alternations are accompanied by a vocalic prothesis when they concern the first vowel of

the word: *lev* – *lv-a* ‘lion Nsg, Gsg’. The Belarusian case is discussed at greater length in Scheer (2009b; 2012, §285).

Finally, the French situation is examined by Pagliano (2003, 815ff). At first sight, French seems to counter-indicate (6): word-initial clusters are only #TR, but first vowels of words freely alternate.⁹ This is true utterance-internally (*la semaine* ‘the week’ may be pronounced either [səmən] or [smən]) as well as utterance-initially (the first word of *reprends du gâteau!* ‘take more cake!’ may be pronounced [χprā] by some speakers). Although the native vocabulary of French features exclusively TR-only items, it thus appears that the language does not ban #RT clusters if produced by phonological computation. This, concludes Pagliano, may indicate that French is currently losing the initial CV (or has lost it already): the lexical TR-only stock will remain for some time. It witnesses a previous state of the grammar, rather than its current shape.

Note that there is indeed a substantial difference between (6b) and (6a, c): the latter look at a result of phonological computation (strength of initial consonants, syncope of first vowels), while the latter examines a *lexical* property of morphemes. There is a link between computation and the shape of lexical entries – lexicon optimization –, but the modification of an entire lexicon is certainly not something that is expected to be carried out over night. Hence a delay in this area is what is awaited – unlike for the other two patterns, which depend on computation and change when computation changes (see Scheer 2012, §294 on this issue).

Also consider that anything-goes languages are typically created on the grounds of TR-only languages when first vowels of words are lost (Scheer 2012, §§328f): this is precisely the case of Slavic (where yers faded away) and modern Occidental varieties of Arabic (such as Moroccan Arabic, where short vowels were lost). Both Common Slavic and Classical Arabic were TR-only.

The theory makes two more predictions, which however cannot be fleshed out for lack of space: (1) the aforementioned fact that in anything-goes languages, missing #RTs are accidental, not systematic gaps (see note 6); (2) languages that allow for initial clusters may be either TR-only or anything goes, but RT-only languages cannot exist. The latter prediction is the spine of Lowenstamm’s (1999) article, and obviously correct.

5. The initial CV is not recorded in the lexical entry of its host

5.1 The initial CV must not be a sleeper

In Lowenstamm’s (1999, 2002) original conception (see also Balogné-Bércecs 2005), the initial CV is part of the lexical recording of words and will be present in the

phonology any time such a word appears. In this perspective, rather than being absent, the initial CV is invisible or inert in languages where no effect is observed at the left edge. The initial CV is thus present at the beginning of the word in all languages, but it is only activated in some through a parameter setting.

There is reason to believe that things do not play out this way.¹⁰ The first thing to point out is certainly that words are not stored in the lexicon; therefore the initial CV cannot be a lexical property of words. It could be a lexical property of roots for example, but we are talking about effects that occur at the left edge of *words* (which may include prefixes), not of roots.

The aforementioned sleepers are a further issue: it was shown in Section 2 that the groundbreaking import of the initial CV is the fact that it is a phonologically meaningful object whose sheer *presence* has a calculable effect. This is how the Direct Effect is produced, in contrast to diacritic sleepers such as # or ω that have no effect by themselves: they only passively sit in representations and wait to be appealed to by some phonological instruction. Were the initial CV always present but only activated in some cases, it would be demoted to the status of a diacritic sleeper.

A related issue showing that the initial CV cannot be recorded in the lexical entry of words is a better understanding of the unit that the initial CV is actually initial of: its effects are observed word-initially in some languages, but they may also occur at the left edge of larger chunks. In languages where phonology applies across word boundaries, initial effects are typically observed at the beginning of utterances and when words are quoted in isolation. Two examples of this pattern are middle Italian dialects (including Corsican, Scheer 2009a; 2012, §270) and Belarusian (Scheer 2009b; 2012, §285). Consider for example the aforementioned prothesis in the latter language. *Lev* 'lion' appears without prothesis when the root vowel is pronounced. When it is not, i.e. in presence of vowel-initial suffixes such as genitive sg. *-a*, in utterance-internal position the presence of the prothesis depends on whether the preceding word is consonant- or vowel-final: compare *brat i-lva* 'the brother of the lion' and *śastra lv-a* 'the sister of the lion'. In utterance-initial position and when the word is quoted in isolation, no calculus can be made with respect to a preceding word – but the prothesis also appears. The reason why a vowel is inserted after consonant-final words is that there are two empty nuclei in a row: /...Cø løv-a/ → [...Ci-løv-a]. Therefore an empty nucleus, i.e. the initial CV, must also be present to the left of utterance-initial /løv-a/. Hence in Belarusian the initial CV is utterance-, not word-initial. Note that it must not be present word-initially: otherwise a prothetic vowel would always appear to the left of /løv-a/ irrespectively of whether the preceding word is consonant- or vowel-final.

This pattern could be adapted to a perspective that uses sleepers: all words in Belarusian come with an initial CV, but only a subset of these CVs is activated,

i.e. those that are utterance-initial and used in quotation. Rather than parametric as in the general case discussed above, the decision to activate the initial CV is grammatical in this scenario. Sleeping CV units, then, would have exactly the status of units of the prosodic hierarchy. All seven layers of these are also always there, but only a subset of them will actually play a role in a particular language, or in a particular process thereof (Nespor and Vogel 1986, 11).

5.2 (Non-) privacy: an echo of SPE

The issue regarding sleepers is also related to a fundamental design property of interface theories that is never discussed in the literature (and cannot be developed here in relevant detail for lack of space): privacy (Scheer 2011, §§383, 400; 2012, §78). It is an undisputed fact that only a small subset of the morpho-syntactic information available is actually relevant for phonology: most of it has no phonological effect at all. This basic observation was first formulated by Chomsky *et al.* (1956), who conclude that phonology is underfed by translational activity (mapping) which transforms morpho-syntactic information into items that are inserted into the phonological string. That is, phonologically irrelevant information has never been translated and is thus absent from phonology.

In spite of this basic observation, SPE reverts back to non-privative translation. Like most other architectural properties of SPE, this stance then abides in all subsequent interface theories.¹¹ That is, full morpho-syntactic information is shipped to phonology regardless of whether it will be used or not. Phonology then appeals to whatever is relevant for its computation; finally, irrelevant information, which is present in form of SPE-type boundaries or prosodic constituency, is either erased by specific phonological action (as in SPE), or just sits in phonology and remains inert (as in Prosodic Phonology).

On the backdrop of current minimalism, putting computational and representational effort into inserting useless noise into phonology without ever using it is hardly an option. Here and elsewhere (also in other disciplines), a basic principle is certainly that objects can only be present when they produce an effect.

5.3 The initial CV is online-created information

A final reason why the initial CV can hardly be a lexical property of words is that this would predict idiosyncrasy: at least in some languages, random distribution across the lexicon would be expected. This pattern, however, appears to be absent from the record. Either the initial CV produces an effect on all items of a specific morpho-syntactically defined category (words, utterances, etc.), or on none.

This shows that the initial CV represents online-created information that comes into being only through morpho-syntactic computation and hence is absent from the lexicon.

6. The initial and other CVs

6.1 Carriers of morpho-syntactic information reduce to syllabic space

When looked at from a broader interface perspective, the initial CV turns out to be but a special case of a larger pattern. If diacritics (such as #s and ωs) are ruled out and if it is recognized that melody (i.e. items below the skeleton) never carry morpho-syntactic information either (Bendjaballah 2012, 6; Bendjaballah and Haiden 2013; Scheer 2011, §660; 2012, §124), the output of translation of morpho-syntactic information reduces to syllabic space (which incarnates into specific representational units according to the phonological theory chosen: skeletal slots, moras, onsets, rhymes etc.). This is a central piece of Direct Interface (Scheer 2012), which is called direct because there are no intermediate categories (#s, ωs etc.) between morpho-syntactic information and phonological objects. Carriers of morpho-syntactic information must be truly phonological items, i.e. which exist in phonology also in absence of any interface event (#s and ωs do not).

If the only object that is ever inserted into the phonological string in order to carry boundary (i.e. non-morphemic) information is syllabic space, we expect to see its effect all over the place when the course of phonology is altered by morpho-syntax. That is, in a theory where syllabic space incarnates as empty CV units, these are expected to be active not only at the beginning of units (word or utterance).

6.2 CV units that have been found to carry morpho-syntactic information

This is indeed what the Government Phonology literature has brought to light. Following up on Guerssel and Lowenstamm (1990), Jean Lowenstamm is at the origin of a research programme that aims to identify the internal structure of templates.¹² The idea is that templates are not just an amorphous set of consonantal and vocalic positions; rather, they have an internal structure. That is, morphological operations do not take place just anywhere in the template: only designated portions are devised for this activity.

Work along these lines includes Bendjaballah (1999 and following), Bendjaballah and Haiden (2003 and following), Lahrouchi (2001 and following), Ségéral (2000), Kihm (2004) and Arbaoui (2010).

Examples of boundary information that is found to be carried by CV units are the negative in Kabyle Berber (Bendjaballah 2001), a verbal marker in Chleuh Berber (Lahrouchi 2001), tense in German strong verbs (Bendjaballah and Haiden 2003) and right-edge demarcation of words (Charette 2003 on Turkish, Luo 2013 on Chinese).

Let us briefly consider the foundational case of Guerssel and Lowenstamm (1990), who argue that Classical Arabic features only one single template (instead of a dozen as in the classical approach). One instrument of this generalization is what they call the derivational syllable, that is, an empty CV unit that occurs between the first and the second CV pair of the unmarked measure I form. On Guerssel and Lowenstamm's analysis, this extra CV unit is always present, but inert (invisible) in measure I. Grammar only makes it accessible in higher, that is derived measures.

The scenario is thus the same as for the initial CV in Lowenstamm (1999): an empty CV unit carries morpho-syntactic and/or semantic information; it is always present in the phonology, but invisible when the corresponding effects are absent.

The only thing that needs to be done in order to make the management of the derivational CV unit privative is to put it under morpho-syntactic control: it realises the specific morpho-syntactic structure that identifies e.g. measure II/III and is therefore inserted into the phonological string via translation only in case this structure exists. Looked at from the phonological side, it is thus present or absent (rather than active or invisible).

This is indeed the direction that further work based on Guerssel and Lowenstamm (1990) has taken: on the analysis of Arbaoui (2010) the four CV units of the template spell out different functional heads of the syntactic tree: C_1V_1 represents V, the derivational syllable C_2V_2 spells out little v, C_3V_3 marks AspP and C_4V_4 is the exponent of AgrP. Guerssel and Lowenstamm's (1990) ready-to-use template is thus entirely pieced together by individual CV units that spell out non-morphemic information.

References

- Arbaoui, Nora. 2010. *Les dix formes de l'arabe classique à l'interface syntaxe/phonologie*. Ph.D. thesis, Université Paris 7.
- Balogné-Bérce, Katalin. 2005. *Strict CV Phonology and the English Cross-Word Puzzle*. Ph.D. dissertation, Eötvös Loránd University, Budapest.
- Bendjaballah, Sabrina. 1999. *Trois figures de la structure interne des gabarits*. Ph.D. dissertation, Université Paris 7.
- Bendjaballah, Sabrina. 2001. "The Negative Preterite in Kabyle Berber." *Folia Linguistica* 34: 185–223.
- Bendjaballah, Sabrina. 2012. *La grammaire des gabarits*. Habilitation à Diriger des Recherches, University Paris 7.

- Bendjaballah, Sabrina, and Martin Haiden. 2003. "Templatic Architecture." *Recherches Linguistiques de Vincennes* 32: 157–168. DOI: 10.4000/rlv.489
- Bendjaballah, Sabrina, and Martin Haiden. 2013. "The Representational Anomalies of Floating Markers: Light Prepositions in Taqbaylit of Chemini." In *Challenges to Linearization*, ed. by Theresa Biberauer, and Ian Roberts, 331–376. Berlin: de Gruyter.
- Brun-Trigaud, Guylaine, and Tobias Scheer. 2010. "Lenition in Branching Onsets in French and in ALF Dialects." In *Development of Language through the Lens of Formal Linguistics*, ed. by Petr Karlík, 15–28. Munich: Lincom.
- Charette, Monik. 2003. "Empty and Pseudo-empty Categories." In *Living on the Edge. 28 Papers in Honour of Jonathan Kaye*, ed. by Stefan Ploch, 465–479. Berlin, New York: Mouton de Gruyter.
- Chomsky, Noam, Morris Halle, and Fred Lukoff. 1956. "On Accent and Juncture in English." In *For Roman Jakobson. Essays on the Occasion of his Sixtieth Birthday*, ed. by Morris Halle, Horace Lunt, Hugh McLean, and Cornelis van Schooneveld, 65–80. The Hague: Mouton.
- Guerssel, Mohand, and Jean Lowenstamm. 1990. "The Derivational Morphology of the Classical Arabic Verb." Ms., UQAM and University Paris 7.
- Kihm, Alain. 2004. "Les pluriels internes arabes: un essai d'anatomie." *Bulletin de la Société de linguistique de Paris* 99: 311–361. DOI: 10.2143/BSL.99.1.541917
- Kijak, Artur. 2005. *Polish and English Complex Consonantal Onsets: A Contrastive Analysis within the Government Phonology Framework*. Ph.D. dissertation, University of Katowice.
- Lahrouchi, Mohamed. 2001. *Aspects morpho-phonologiques de la dérivation verbale en berbère (parler chleuh d'Agadir)*. Ph.D. dissertation, Université Paris 7.
- Lowenstamm, Jean. 1996. "CV as the Only Syllable Type." In *Current Trends in Phonology. Models and Methods*, vol. 2, ed. by Jacques Durand, and Bernard Laks, 419–441. Salford, Manchester: ESRI. Available at <http://sites.unice.fr/scheer>.
- Lowenstamm, Jean. 1999. "The Beginning of the Word." In *Phonologica 1996*, ed. by John Rennison, and Klaus Kühnhammer, 153–166. La Hague: Holland Academic Graphics. Available at <http://sites.unice.fr/scheer>.
- Lowenstamm, Jean. 2002. "A Note on the Segmental Identification of Templatic Sites." Ms., University Paris 7.
- Luo, Xiaoliang. 2013. *Vers une phonologie gabaritique du chinois. Analyse des interactions entre syllabe, ton et processus morphologiques*. Ph.D. thesis, Université d'Orléans.
- Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Pagliano, Claudine. 2003. *Lépenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie*. Ph.D. dissertation, Université de Nice.
- Pak, Majorie. 2008. *The Postsyntactic Derivation and its Phonological Reflexes*. Ph.D. dissertation, University of Pennsylvania.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Linguistic Inquiry Monograph 16. Cambridge, Mass.: MIT Press.
- Samuels, Bridget. 2009. *The Structure of Phonological Theory*. Ph.D. dissertation, Harvard University.
- Sanoudaki, Eirini. 2010. "Towards a Typology of Word Initial Consonant Clusters: Evidence from the Acquisition of Greek." *Journal of Greek Linguistics* 10: 1–41. Available at <http://sites.unice.fr/scheer>. DOI: 10.1163/156658410X495826
- Scheer, Tobias. 1999. "A Theory of Consonantal Interaction." *Folia Linguistica* 32: 201–237. Available at <http://sites.unice.fr/scheer>.
- Scheer, Tobias. 2000. *De la Localité, de la Morphologie et de la Phonologie en Phonologie*. Habilitation thesis, Université de Nice. Available at <http://sites.unice.fr/scheer>.
- Scheer, Tobias. 2004. *A Lateral Theory of Phonology. Vol.1: What is CVCV, and why Should it Be?* Berlin: Mouton de Gruyter. DOI: 10.1515/9783110908336
- Scheer, Tobias. 2007. "On the Status of Word-Initial Clusters in Slavic (And Elsewhere)." In *Annual Workshop on Formal Approaches to Slavic Linguistics. The Toronto Meeting 2006*, ed. by Richard Compton, Magdalena Goledzinowska, and Ulyana Savchenko, 346–364. Ann Arbor: Michigan Slavic Publications. Available at <http://sites.unice.fr/scheer>.
- Scheer, Tobias. 2008. "Why the Prosodic Hierarchy is a Diacritic and Why the Interface Must Be Direct." In *Sounds of Silence: Empty Elements in Syntax and Phonology*, ed. by Jutta Hartmann, Veronika Hegedüs, and Henk van Riemsdijk, 145–192. Amsterdam: Elsevier. Available at <http://sites.unice.fr/scheer>.
- Scheer, Tobias. 2009a. "External Sandhi: What the Initial CV is Initial of." *Studi e Saggi Linguistici* 47: 43–82. Available at <http://sites.unice.fr/scheer>.
- Scheer, Tobias. 2009b. "Representational and Procedural Sandhi Killers: Diagnostics, Distribution, Behaviour." In *Czech in Formal Grammar*, ed. by Mojmír Dočekal, and Markéta Ziková, 155–174. München: Lincom. Available at <http://sites.unice.fr/scheer>.
- Scheer, Tobias. 2011. *A Guide to Morphosyntax-Phonology Interface Theories. How Extra-Phonological Information is Treated in Phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton de Gruyter.
- Scheer, Tobias. 2012. *Direct Interface and One-Channel Translation. A Non-Diacritic Theory of the Morphosyntax-Phonology Interface. Vol.2 of a Lateral Theory of Phonology*. Berlin: de Gruyter.
- Scheer, Tobias, and Markéta Ziková. 2010. "The Coda Mirror v2." *Acta Linguistica Hungarica* 57 (4): 411–431. DOI: 10.1556/ALing.57.2010.4.4
- Ségéral, Philippe. 2000. "Théorie de l'apophonie et organisation des schèmes en sémitique." In *Research in Afroasiatic Grammar*, ed. by Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 263–299. Amsterdam: John Benjamins. Available at <http://sites.unice.fr/scheer>.
- Ségéral, Philippe, and Tobias Scheer. 2001. "La Coda-Miroir." *Bulletin de la Société de Linguistique de Paris* 96: 107–152. Available at <http://sites.unice.fr/scheer>. DOI: 10.2143/BSL.96.1.503739
- Ségéral, Philippe, and Tobias Scheer. 2008a. "Positional Factors in Lenition and Fortition." In *Lenition and Fortition*, ed. by Joaquim Brandão de Carvalho, Tobias Scheer, and Philippe Ségéral, 131–172. Berlin: Mouton de Gruyter. Available at <http://sites.unice.fr/scheer>.
- Ségéral, Philippe, and Tobias Scheer. 2008b. "The Coda Mirror, Stress and Positional Parameters." In *Lenition and Fortition*, ed. by Joaquim Brandão de Carvalho, Tobias Scheer, and Philippe Ségéral, 483–518. Berlin: Mouton de Gruyter. Available at <http://sites.unice.fr/scheer>.
- Seigneur-Froli, Delphine. 2006. *Le Statut phonologique du début de mot grec. Lenitions consonantiques et libertés phonotactiques initiales dans la diachronie de la langue commune et dans le dialecte de Lesbos*. Ph.D. dissertation, University of Nice. Available at <http://sites.unice.fr/scheer>.

Notes

1. The article is based on Direct Interface as exposed in Scheer (2012). Space restrictions preclude discussion of a number of relevant issues, the most obvious being that only the representational aspect of the interface is covered. There is also a derivational aspect (cyclic derivation, today phase theory) which remains unmentioned. See also note 7.

2. A reviewer points out that the division of the linear string into phonologically relevant chunks may be done by diacritics, which then are able to produce predictions. He quotes an example whereby prefix-stem constructions are prosodified as [_ω [_ω prefix] [_ω stem]]: if there is then a bimoraic minimality requirement observed on roots, a prediction is made to the effect that prefixes will also be minimally bimoraic. This is because both the prefix and the root chunk have the same label: an omega. In English, this prediction is borne out: word-level prefixes like *un-* are obligatorily bimoraic and full-vowelled. This is all correct, but does not speak to the issue: the effect achieved is due to chunking of the linear string, not to any intrinsic properties of omegas. If the labels of the chunks were bananas, the bimoraic requirement on banana-chunks would produce exactly the same result. Omegas have no phonological property at all and hence are unable to make any phonological prediction.
3. Note that rule (1)a says that vowels are deleted *only* when they occur in word-initial clusters. Of course there are languages where vowels are deleted in this context (e.g. Czech *pes* – *ps-a* ‘dog Nsg, Gsg’), but they will then also be deleted elsewhere (Czech *loket* – *lokt-e* ‘elbow Nsg, Gsg’). In the same way, rule (1)b inserts a vowel into clusters *only* when they are word-initial. Non-initial clusters remain unrepaired.
4. In this article, T is shorthand for obstruents, R for sonorants.
5. This is the representation of TR clusters that appears in Brun-Trigaud and Scheer (2010) and Scheer (2000), where its local character is motivated by the parallel with locality restrictions on movement in syntax (Relativized Minimality, e.g. Rizzi 1990). Earlier representations of TR clusters in CVCV were not local in kind.
6. The fact that really *anything* goes in these languages, not just a subset of logically possible clusters (i.e. grammar does not object to any initial cluster), is discussed at greater length in Seigneur-Froli (2006), Sanoudaki (2010) and Scheer (2012, §313).
7. The global situation is in fact more intricate. A basic insight of Lexical Phonology is that the impact of a given morpho-syntactic division may be process-specific, i.e. block one phonological process, but not another (in the same language). This further complication cannot be covered in this article. It is discussed in Scheer (2011, §§811, 823; 2012, §294).
8. Some caution is in order regarding the interpretation of the relevant data: Ségéral and Scheer (2008a, 150ff) discuss this issue.
9. In case they are schwa, and if they are not preceded by a TR cluster: for instance, schwa may not be dropped in *la gr[ə]nouille* ‘frog’ (see Scheer 2004, §90).
10. Scheer (2000, 273ff) provides some early discussion of the issue.
11. In the trail of SPE, Lowenstamm (1999, 164) also argues that the distribution of the initial CV is restricted to major categories, i.e. nouns, verbs and adjectives. As far as I can see, no case has been documented where an initial effect is observed on nouns, verbs and adjectives but not, say, on prepositions, adverbs and the like.
12. The programme is fleshed out in number 32 of the journal *Recherches Linguistiques de Vincennes* (2003), which is guest-edited by J. Lowenstamm and contains a number of relevant articles.

PART 5

Selected Issues in Afro-Asiatic (Morpho-) Syntax and Semantics