SYNC
2016

Program
& Abstracts

18th Annual Meeting of Linguists from
Stony Brook, Yale, NYU, and CUNY GC

December 3rd, 2016, at The Graduate Center
City University of New York
CONTENTS

Program........................................................................................................................................ 4
Invited Speaker.................................................................................................................................. 6
Abstracts........................................................................................................................................... 7

Some notes on “infinitival perfects” in Appalachian English
Christina Tortora............................................................................................................................. 8

Sociolinguistically-deduced sound change in Zrawa variety of Tunisian Tamazight: from interdentals to stops
Wafa Bahri (CUNY GC)................................................................................................................... 9

NPIs and modals: two reasons “just” and “only” are not the same
Taletha Callahan-Kanik (CUNY GC)................................................................................................ 11

Phrase-final so: an Inquisitive Semantics approach
Mary Robinson (NYU)....................................................................................................................... 13

Variation in stem-final consonant clusters in Korean nominal inflection
Ji Yea Kim (Stony Brook)................................................................................................................ 16

A Model-Theoretic Approach to Minimalist Syntax
Yiding Hao (Yale)............................................................................................................................ 18

Unifying the that-t effect and the do-t effect
Hwichan Oh (Stony Brook)............................................................................................................... 20

Oblique causer constructions in Lezgian
Philip Shushurin (NYU)................................................................................................................ 22

Dependencies Between Adverbs and Sentential Clitics: A Crosslinguistic Perspective
Russell Tanenbaum (Stony Brook).................................................................................................. 24

Homeostatic Reinforcement Learning for Harmonic Grammars
Jon Rawski (Stony Brook)............................................................................................................... 26

TextPros: Prosody-text aligner for unlabeled speech
Hussein Ghaly (CUNY GC)............................................................................................................ 28
Tone analysis in Tai Khambi: computational models in language documentation
Rikker Dockum (Yale) ........................................................................................................ 30

Local services: food & printing ......................................................................................... 32
Area map ............................................................................................................................ 33
Participants .......................................................................................................................... 34
Acknowledgments .............................................................................................................. 35
SYNC 2016 Program
Saturday, December 3 - Room 5414, The Graduate Center, CUNY

9:00 - 10:00  Check-in
            room 5414, coffee and light breakfast provided

10:00 - 10:05 Opening remarks

10:05 - 12:05 Session 1: Semantics/Sociolinguistics - Chair: Hussein Ghaly
              10:05 - 10:35 Sociolinguistically-deduced sound change in Zrawa variety of Tunisian
              Tamazight: from interdentals to stops
              Wafa Bahr (CUNY GC)
              10:35 - 11:05 NPIs and modals: two reasons “just” and “only” are not the same
              Tally Callahan-Kanik (CUNY GC)
              11:05 - 11:35 Phrase-final so: an Inquisitive Semantics approach
              Mary Robinson (NYU)
              11:35 - 12:05 Variation in stem-final consonant clusters in Korean nominal inflection
              Ji Yea Kim (Stony Brook)

12:05 - 1:30 Lunch break

1:30 - 3:30 Session 2: Syntax – Chair: Tally Callahan
            1:30 - 2:00 A Model-Theoretic Approach to Minimalist Syntax
            Yiding Hao (Yale)
            2:00 - 2:30 Unifying the that-t effect and the do-t effect
            Hwichan Oh (Stony Brook)
            2:30 - 3:00 Oblique causer constructions in Lezgian
            Philip Shushurin (NYU)
            3:00 - 3:30 Dependencies Between Adverbs and Sentential Clitics: A Crosslinguistic
            Perspective
            Russell Tanenbaum (Stony Brook)
3:30 - 4:00  Coffee break
coffee provided

4:00 - 5:30  Session 3: Phonology – Chair: Wafa Bahri

4:00 - 4:30  Homeostatic Reinforcement Learning for Harmonic Grammars
Jon Rawski (Stony Brook)

4:30 - 5:00  TextPros: Prosody-text aligner for unlabeled speech
Hussein Ghaly (CUNY GC)

5:00 - 5:30  Tone analysis in Tai Khamti: computational models in language
documentation
Rikker Dockum (Yale)

5:30 - 6:30  Invited talk

Some notes on “infinitival perfects” in Appalachian English
Christina Tortora (CUNY GC)

6:30 - 9:00  Reception
food and beverages provided in the Linguistics Lounge, room 7400
Christina Tortora is a Professor of Linguistics at the College of Staten Island and at the Graduate Center CUNY. She defended her dissertation entitled "The Syntax and Semantics of the Weak Locative" at the department of Linguistics at the University of Delaware. Professor Tortora has published various books, articles in linguistics journals, and chapters in books (both in English and Italian) on various topics such as the syntax of Italian dialects, existential constructions, the nature of so-called pleonastic elements in grammar (also known as “expletives”), possessive constructions, prepositions and prepositional syntax, aspect, resultative predication, clause structure, auxiliary selection, and agreement.

She has taught at numerous universities including University of Michigan at Ann Arbor, University of Padua, Italy and State University of New York at Stony Brook, and has given many invited lectures and talks both nationally and internationally.

In 2001, she held a National Endowment for the Humanities (NEH) Fellowship to do research for a grammar of Borgomanerese, and has been awarded several NSF grants to support her research and conference organization. In 2006, she was the humanities recipient of the Feliks Gross Endowment Award for Outstanding Scholarly Achievement by a junior faculty member, awarded by the CUNY Academy for the Humanities and Sciences. She is the Principal Investigator on two projects: AAPCAppE - The Audio-Aligned and Parsed Corpus of Appalachian English with the target of creating a 1 million word corpus of Appalachian English as well as CONYCE (Corpus of New York City English) along with Cece Cutler, Bill Haddican and Michael Newman, for which they recently received a National Science Foundation grant.
SYNC 2016
ABSTRACTS
Some notes on “infinitival perfects” in Appalachian English

Christina Tortora
CUNY (College of Staten Island and The Graduate Center)

In this talk (which includes collaborative work with Beatrice Santorini and Greg Johnson), I explore data from a series of related constructions in Appalachian English, with an eye towards gaining a more nuanced understanding of an English tense often referred to in the literature as the “infinitival perfect.” The data come from the Audio-Aligned and Parsed Corpus of Appalachian English (the AAPCAppE; Tortora et al. to appear).

Consider the examples in (1) from the AAPCAppE, which help to give an initial sense of the phenomenon under study:

(1) a. I don’t know what I would done if it wasn’t for her.
   b. That was supposed to been a rare seed.

If we think of these examples in the context of other Englishes, it seems that there is a missing infinitival auxiliary have in (1a) and (1b); that is, these examples appear at first glance to be the equivalent of (2a) and (2b):

(2) a. I don’t know what I would have done if it wasn’t for her.
   b. That was supposed to have been a rare seed.

To better understand the phenomenon in (1), and what it tells us about the so-called infinitival perfect in English more generally, I present a preliminary quantitative / qualitative study of variation in use of infinitival auxiliary have (typically pronounced [ǝv] or [ǝ]) in perfects embedded under modals and to in the AAPCAppE. The frequency of the presence vs. absence of have in AAPCAppE infinitival perfects shows that there are differences between the modal vs. infinitival-to contexts which have not been previously noted in the literature. The data gleaned from the corpus provide compelling clues suggesting that “infinitival perfects” are not a homogeneous structural or semantic phenomenon. They also illustrate the usefulness of audio-aligned and parsed corpora of vernacular speech as tools for the study of syntax.

References:
The aim of this paper is to discuss variations in the use of the interdental variables /θ/ and /ð/ and their respective counterparts /t/ and /d/ in two micro-varieties of Tunisian Tamazight spoken in the villages of Douiret and Zrawa. While the Douiret variety has no interdental sounds in its inventory as in the Proto-Berber (An Afroasiatic reconstructed proto-language from which all modern Berber/Tamazight languages stem); the Zrawa variety exhibits a process of sound shift, substituting the interdentals with the stops. Research on Tamazight languages demonstrates that similar phonological processes show a reversal change that substitutes synchronic realizations of sounds for historical ones (Saib 1997, Kossmann 1997). In this study, shifting back to the original phonetic realizations is posited to be motivated by sociolinguistic factors such as age, level of education, modernization, and stigmatization.

Data used for this study came primarily from two-recorded word lists by twelve speakers of Tamazight: six speakers from the village of Douiret and six speakers from the village of Zrawa. The ages of the speakers ranged between 18 and 64 years old. All the speakers grew up as bilingual in Tunisian Tamazight and Tunisian Arabic, and had various levels of education. The word lists consisted of thirty five words each with /t/ and /d/ sounds. During the recordings, the twelve speakers did not read from the Tamazight lists, but were given translated lists of the same words in Tunisian Arabic and were asked to give their equivalents in Tamazight. A total of 420 tokens were transcribed using the International Phonetic Alphabet (IPA) and then tabulated and compared.

The analysis confirmed that all the speakers from the village of Douiret show no interdentals in their pronunciation of the word lists. They, instead, produced the stops /t/ and /d/ as in the Proto-Berber (Koussman 2014). Participants from Zrawa, however, displayed clear, socially-conditioned variations in their pronunciations. The directionality reversal showed a great effect of sociolinguistics variables, principally age, education, modernization, and stigmatization. We concluded that Tunisian Tamazight spoken in the village of Zrawa is undergoing a sound change in progress from interdentals to dental stops. While a system-internal explanation could be provided for this change, it remains unsatisfactory without considering the role played by social factors. The findings also propose that the directionality reversal observed in Zrawa variety with regard to these social variables is a result of shifting language ideologies in the village.

The paper is the first to describe a sound change in a Tunisian Tamazight variety and to attribute it to social factors and ideological shift that could be pertained to the nascent linguistic and social Amazigh movements in Tunisia.
References


NPIs and modals: two reasons “just” and “only” are not the same
Taletha Callahan-Kanik

On cursory inspection, the English exclusives just and only appear interchangeable in their parallel non-scalar and scalar usages (a non-scalar usage of just/only enforces a limit of some kind whereas a scalar usage of just/only minimizes the significance of some subsequent content). Many researchers (such as Beaver & Clark 2008) have considered them to be the same in form and function due to this interchangeability in many contexts; however, I argue that they differ, drawing upon data from NPIs and modals.

The two main differences that I have identified between only and just are their ability to license Negative Polarity Items (NPIs) and their interaction with modals. Consider the following examples (capital letters indicate prosodic prominence):

(1) John only ever [DISLIKED Sally]
    Interpretation: At any time in the past, it was not the case that John had any stronger negative feelings for Sally than dislike (e.g. hate)

(2) #John just ever [DISLIKED Sally]
    Note: The # here means that the sentence is unacceptable, or at least does not mean the same thing as (1)
    Interpretation: ???

(3) a. James can only eat BREAD
    Interpretation 1: The one thing John can eat is bread
    Interpretation 2: John is allowed to eat bread and nothing else if he wants to.
    b. James can just eat BREAD
    Interpretation: John is allowed to eat bread and nothing else if he wants to.
    Blocked interpretation: The one thing John can eat is bread

Examples (1) – (3) demonstrate strong differences of interpretation between parallel sentences containing only and just. If they truly have the same structure and semantics as much of the prior literature has assumed, their behavior should reflect that. I’ve observed, however, that only’s behavior is more similar to negation in these two areas. I propose (following Jacobs 1980 and von Fintel & Iatridou 2007) that only is a case of “negative split” while just is a single semantic unit.

My proposed structure and semantics for just are in line with the structure and semantics commonly attributed to only under a classic Horn (1969) and Rooth (1985, 1992) account:

(4) \([\text{just}_C S]\)^w is defined only if (i.e. presupposes that) \([S]\)^w = 1.
    If defined, \([\text{just}_C S]\)^w = 1 \text{ iff } \forall S'(S' \in C) & S \neq S' \rightarrow \([S']\)^w = 0

My proposed semantics and structure for only, however, deviates substantially from this classic view. It is similar to the one proposed by von Fintel & Iatridou: only is composed of negation and a special exceptive (known as QUE, from the ne...que construction found in French). The main difference from their theory is that a different final scope of the NEG and QUE is proposed:
(5) \(\text{[QUE}]^w(x)(P)\) is defined only if (i.e. presupposes that) \(P(x) = 0\) in \(w\).
If defined, \(\text{[QUE}]^w(x)(P) = 1\) iff \(\forall y (y \in D \& y \neq x \Rightarrow P(y) = 1\) in \(w\))

\(D\) = the set of alternatives to the focus associate

\(\text{[QUE]}\) takes two arguments \((x\) and \(P)\). In a world (as designated by the superscript \(w\)),
\(\text{[QUE]}(x)(P)\) is defined only if \(P(x)\) is false. If defined in that world, \(\text{[QUE]}(x)(P)\) is true if and
only if for all \(y\), if \(y\) is in the set of alternatives to the focus associate\((D)\) and \(y\) is not the same as \(x\),
then \(P(y)\) is true in that world. This semantic entry requires that the QUE-phrase undergo
quantifier raising (QR) to be interpretable. In (5), [QUE Beth] is of type \(<e,t,t>\) (the type of a
generalized quantifier). Since generalized quantifiers are only interpretable at clausal nodes, it
must raise. This results in the following structure:

(6) \([\text{[QUE}}^\text{Beth}\][\lambda x.\text{John [NEG likes x]})]

This deviates from von Fintel and Iatridou (2007), since QUE now scopes above NEG.

One crucial difference between my proposed semantics for just and only is that there is no
negation in the structure of just. This lack of negation is the key to the divergence of the two
exclusives regarding NPI-licensing and scope relations when interacting with modals (see below).
Strawson Downward Entailment (von Fintel 1999) has been utilized in the past to account for
only’s ability to license NPIs, but just also fulfills the requirements for SDE, which would anticipate
(2) to be acceptable. With negation as part of only’s composition, we can go back to standard
Downward Entailment as an NPI licensing environment, which would avoid this problem. If
negation is a core part of only, then its negation-like interaction with modals is expected (e.g. the
modals may and can both allow negation/only to scope above or below them), in contrast to how
just interacts with the same modals (e.g. the modals may and can most saliently scope above just):

(7) Negation (NOT)
   a. James may not eat bread. negation > modal OR modal > negation
   b. James can not eat bread.  negation > modal OR modal > negation

(8) Exclusive ONLY
   a. James may only eat [bread]. only > modal OR modal > only
   b. James can only eat [bread].  only > modal OR modal > only

(9) Exclusive JUST
   a. James may just eat [bread]. modal > just (?? just > modal)
   b. James can just eat [bread].  modal > just (?? just > modal)

My current research takes this theory and seeks to gain more supportive evidence from
original experimental data. Such experimental work may prove extremely enlightening, since much
of the work on exclusives has been purely theoretical. The experiment will gather acceptability
judgments from online participants on sentences containing negation, only, and just. All
experimental sentences will have both an NPI and a non-NPI version. I hypothesize that overall the
sentences containing only + NPI will be found more acceptable than the sentences containing just
+ NPI. The results I gain from my experiment will further linguistic understanding of the NPI-
licensing abilities of only and just in English.
Phrase-final so: an Inquisitive Semantics approach
Mary Robinson, New York University

**Introduction:** Theoretical frameworks should not only account for empirical data they analyze, but also new empirical phenomena that appear in the course of natural language change. In this paper, I test whether the conversational maxims originally proposed by Grice (1975), and updated by Groenendijk & Roelofson (2009) in the Inquisitive Semantics framework are able to do so. The test case is a discourse particle in English that I have named Phrase Final so (PFS). I will show that the cooperative way in which speakers use this particle to exchange information represents a novel way that speakers make use of sets of alternatives to resolve issues. Speakers make one proposition more salient than others, and close the utterance with PFS spoken with a short duration and falling intonation. This PFS communicates that there is only one alternative that can settle the issue at hand, the one containing the informative proposition which immediately precedes so. In Inquisitive Pragmatics terms, PFS turns a compliant answer into an optimally compliant answer.

**Phrase-final so:** An apparently undescribed recent innovation in English is the use of *so* to end an utterance as a response to a polar question or proposition.

(1) A: Do you want to get dinner?  
B: I teach in the morning, so.  
“Because of my teaching obligation, I cannot stay out late (i.e., go to dinner).”

(2) A: She’s very polite!  
B: She’s Canadian, so.  
“Yes, she is polite because she is Canadian, and all Canadians are polite.”

As the paraphrases show, PFS relates the new information to the original proposition or question. Informally, PFS appears to be a particle that is elliptical and anaphoric to one of the alternatives presented by the questions or statements that speaker A makes in examples (1) and (2). In other words, PFS indicates which of the contextually salient alternatives has been selected.

**Inquisitive Semantics:** This framework encompasses a notion of meaning that concerns information exchange potential, with a conversation between two participants revolving around raising and resolving issues (Ciardelli et al 2013). Those notions are defined as follows:

*Issues* (Ciardelli et al 2013: 461): Let $s$ be an information state, and $I$ a non-empty set of enhancements of $s$. Then we say that $I$ is an issue over $s$ if and only if:

a. $I$ is downward closed: if $t \in I$ and $t' \subset t$ then also $t' \in I$.

b. $I$ forms a cover of $s$: $\cup I = s$.

*Settling an issue* (Ciardelli et al 2013: 461): Let $s$ be an information state, $t$ an enhancement of $s$, and $I$ an issue over $s$. Then we say that $t$ settles $I$ if and only if $t \in I$.

In other words, within Inquisitive Semantics, issues are requests for information to locate the actual world within a set of possible worlds. Settling an issue involves enhancing the discourse context to pick out the actual world, or at least narrow down the possibilities. Note that the enhancements of information states are downward closed, i.e. that every specific enhancement of the discourse context must be contained in a more general enhancement.

**Analysis:** A discourse context $c$ is non-empty and downward closed, containing all the information that has been established and the issues raised. The speakers will continue to narrow this set of possible worlds, never reintroducing a world that has already been eliminated. In the case of example (1), A first raises an issue, asking if B wants to get dinner. The set of
enhancements introduced by ?GetDinner divide the possible worlds into two alternatives: those in which A and B get dinner, and those in which they do not (see Fig. 1a).

Now that A has raised an issue, B must give a compliant response. The notion of compliance is derived from Grice’s (1975) Maxim of Quality, and states that the response $\varphi$ to an inquisitive proposition $\psi$ is optimally compliant if it picks out exactly one alternative raised by $\psi$. However, not all responses are optimally compliant. The response $\varphi$ is still compliant with $\psi$ if $\varphi$ partially settles the issue raised by $\psi$ (Groenendijk & Roelofson 2009).

Without PFS, B’s response partially resolves the issue by reducing the number of possible worlds that could be the actual world. However, the answer is not optimally compliant because A’s request for information is not settled. As Figure 1b shows, there are still two alternatives still at play: a) yes, get dinner & teach in the morning (the 11 worlds); and b) no, do not get dinner & teach in the morning (the 01 worlds). B has updated the discourse context and eliminated two alternatives, but this the most amount of information that B can give without the PFS.

When B ends the utterance with PFS, the answer is optimally compliant because it picks out only one alternative: that in which B does not go to dinner and does teach in the morning (01). The inclusion of PFS has the effect of highlighting the proposition Teach in the morning (i.e. making it the most salient proposition; see Roelofsen & Van Gool 2010). This narrows down the sets of possible worlds to one, as in Figure 1c. This set contains the highlighted proposition Teach in the morning, but not Go to dinner.

Figure 1: (a) the polar question ?GetDinner, (b) after the enhancement of B’s informative proposition (before PFS), and (c) the effect of following the proposition with PFS.

Discussion: The Inquisitive Semantics framework allows a straightforward analysis of new phenomena that occur in natural language. The discourse particle PFS turns a compliant answer into an optimally compliant answer because it picks out the relevant proposition within the sets of alternatives and asserts that the set chosen will contain only this proposition and no other contextually salient ones. Interestingly, when both propositions are mentioned in the answer, it is the conjunct preceding PFS that determines which alternative is chosen. Consider the two possible responses to A’s invitation to dinner:

(3) a. I teach in the morning, but I want to try the new restaurant around the corner, so.

b. I want to try the new restaurant around the corner, but I teach in the morning, so.

Example (3) shows that PFS is not the same as sentence final particles (SFP). SFPs (e.g. eh or huh) are merged in the Speech Act layer of the syntax, whereas PFS is merged in the syntactic derivation as a complementizer with an elided TP complement: \[CP so [\{TP \text{I won’t go}\}].\] Therefore, PFS highlights only the preceding clause, causing the contrast between (3a) and (3b). As a result, PFS can make more fine-grained distinctions in choosing alternatives than SFPs can.

Variation in stem-final consonant clusters in Korean nominal inflection

Ji Yea Kim

Introduction. Consonant clusters are prohibited in Korean, undergoing consonant cluster simplification in isolation (e.g. /kaps/ [kap] ‘price’). Underlying clusters in coda are, however, preserved in surface representations when a vowel-initial suffix is attached. They are further resyllabified as a repair strategy to conform to the (C)V.C.CV structure (e.g. /kaps-i/ [kap.si] ‘price-NOM’). Such preservation is considered as being standard, which is phonologically predictable. It is also notable that one of the two consonants in stem-final clusters can optionally be deleted when combining with a vowel-initial suffix (e.g. /kaps-i/ [kap.si]–[ka.pi] ‘price- NOM’). This is nonstandard, which has been shown in previous research as variation that the younger generation of Korean speakers has recently exhibited (e.g. Kenstowicz 1996, Kim 2005). However, little is known about linguistic and social factors that affect younger speakers’ variation. This study thus aims to provide an extensive analysis as to whether and to what extent consonant cluster types, suffix types, or their interaction serve as a significant factor in selecting between standard forms and nonstandard forms for inflected nouns in Korean. In addition, speakers’ dialects (e.g. Seoul and South Kyungsang) and gender are examined.

Experiment. A production experiment was conducted with a total of 10 native speakers of Korean (mean age: 23; range: 19-28). There were five male speakers and five female speakers. Five were Seoul Korean speakers while the other five were South Kyungsang Korean speakers. For speech materials, 27 monosyllabic native Korean nouns with a coda cluster and a vowel-initial suffix attached were used: the four clusters /ps/, /ks/ (i.e. ‘stop + fricative’), /lk/, and /ls/ (i.e. ‘liquid + stop/fricative’) with the three suffixes -i (NOM), -il (ACC), and -e (LOC/DAT) (Table 1). Items were printed on a sheet of paper in a random order, embedded in the medial position of present tense declarative sentences (e.g. ćo ki-pota kaps-i pis’a-jo ‘(Its) price is expensive than there [that].’). There were also 12 fillers, which included four other coda clusters in verb stems with three vowel-initial verbal suffixes. Each participant was asked to read 39 sentences (= 27 targets + 12 fillers) with three repetitions in a natural tone and at a normal speed. In this way, a total of 1,170 sentences were recorded (= 39 targets and fillers x 3 repetitions x 10 speakers). Out of the 1,170 tokens, 810 tokens were examined (= 9 test items x 3 suffixes x 3 repetitions x 10 speakers), excluding the 360 fillers (= 12 fillers x 3 repetitions x 10 speakers). The overall number of standard and nonstandard forms was judged and counted by the researcher, following Jun and Lee (2007).

Results and discussion. Results showed that there was a significant main effect for clusters (p<.001) but not for suffixes (p>.05) and their interaction (p>.05) on younger speakers’ selection between standard forms and nonstandard forms in Korean nominal inflection (Figure 1). Speakers were likely to preserve both consonants for the clusters of ‘stop + fricative’ (Percentages of preservation: 98.89% for /ps/; 80.74% for /ks/). In contrast, speakers tended to delete one of the two consonants for the clusters of ‘liquid + stop/fricative’ (Percentages of preservation: 10.83% for /lk/; 8.89% for /ls/). In most cases, the liquid was more subject to deletion whose motivation from a constraint-based perspective is that the constraint [-continuant, -coronal] dominates [+lateral] syllable-finally in Korean (Ahn 1992).

What is more intriguing is that /l/-deletion was found across dialects in the present study and is similar to the pattern reported for Seoul speakers’ production approximately thirty years ago: in the ‘liquid + /k, p, pʰ/’ clusters, /l/ was always preserved in the Kyungsang dialect (Whitman 1985) while C2 was maintained in the Seoul dialect (Yu Cho 1988). In other words, there was no significant difference according to speakers’ dialects in present-day Korean: Seoul speakers (M=1.43, SD=.50) and South Kyungsang speakers (M=1.49, SD=.50); p>.05. This indicates that South Kyungsang speakers have assimilated their speech to Seoul speakers’ production. Similarly, Jun and Lee (2007) mentioned that their participants who were North Kyungsang Korean speakers in their 20’s had been exposed to standard Seoul Korean via the mass media, consequently showing a similar pattern with Seoul speakers’ speech.

In terms of gender, male speakers used more nonstandard forms (M=1.40, SD=.49) whereas female speakers favored standard forms (M=1.52, SD=.50) with a significant difference between them (p<.001). This is in line with Labov’s (1990) Principle I in that, for stable sociolinguistic variables, men use more nonstandard forms than women do. Female speakers in the current study tended to prefer standard forms by preserving clusters, which are learned in school and used in the mass media. This is
presumably to conform to conservative social pressures and expectations that apply to women in Korea. In sum, this study shows that some linguistic and social factors play a significant role in selecting variants among inflected forms of nouns containing stem-final consonant clusters in contemporary Korean.

**Table and figure**

Table 1. List of 27 nouns with four stem-final consonant clusters and three vowel-initial suffixes (based on the Great Standard Korean Dictionary provided by the National Institute of the Korean Language)

<table>
<thead>
<tr>
<th>Manner of articulation</th>
<th>Clusters</th>
<th>Test stems</th>
<th>Nominative suffix -i</th>
<th>Accusative suffix -ɨl</th>
<th>Locative/dative suffix -e</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop + fricative</td>
<td>ps</td>
<td>kaps</td>
<td>kaps-i</td>
<td>kaps-ɨl</td>
<td>kaps-e</td>
<td>'price'</td>
</tr>
<tr>
<td></td>
<td>ks</td>
<td>saks</td>
<td>saks-i</td>
<td>saks-ɨl</td>
<td>saks-e</td>
<td>'wage'</td>
</tr>
<tr>
<td></td>
<td>moks</td>
<td>moks-i</td>
<td>moks-ɨl</td>
<td>moks-e</td>
<td></td>
<td>'share'</td>
</tr>
<tr>
<td></td>
<td>REAKS</td>
<td>REAKS-i</td>
<td>REAKS-ɨl</td>
<td>REAKS-e</td>
<td></td>
<td>'soul'</td>
</tr>
<tr>
<td>Liquid + stop</td>
<td>lk</td>
<td>talk</td>
<td>talk-ɨl</td>
<td>talk-e</td>
<td></td>
<td>'chicken'</td>
</tr>
<tr>
<td></td>
<td>salk</td>
<td>salk-ɨl</td>
<td>salk-e</td>
<td></td>
<td></td>
<td>'leopard cat'</td>
</tr>
<tr>
<td></td>
<td>ĉʰilk</td>
<td>ĉʰilk-ɨl</td>
<td>ĉʰilk-e</td>
<td></td>
<td></td>
<td>'arrowroot'</td>
</tr>
<tr>
<td></td>
<td>hilk</td>
<td>hilk-ɨl</td>
<td>hilk-e</td>
<td></td>
<td></td>
<td>'soil'</td>
</tr>
<tr>
<td>Liquid + fricative</td>
<td>ls</td>
<td>kols</td>
<td>kols-ɨl</td>
<td>kols-e</td>
<td></td>
<td>'way'</td>
</tr>
</tbody>
</table>

Figure 1. Number of stem-final consonants produced with respect to four clusters and three suffixes

**References**


Unifying the *that*-t effect and the *do*-t effect

Hwichan Oh

**Introduction:** Pesetsky and Torrego (2001, henceforth P&T) analyze the *that*-t effect and the ban on T-to-C movement in *wh*-subject question without emphatic meaning (the *do*-t effect in Pesetsky (2015)) in a unified way, as an economy violation. Their insight is that the occurrence of an overt C element is forbidden in both cases (Koopman 1983).

(1) Who do you think (*that) t left? (the *that*-t effect)
(2) Who (*did) t put an apple on the table? (the *do*-t effect)

To account for these sentences, P & T make some assumptions; (i) nominative case is uT on D, (ii) the head C has uT, (iii) the complementizer *that* is a result of T-to-C movement and (iv) either T-to-C movement or the movement of a subject in [Spec, TP] to [Spec, CP] can check uT on C. When a subject is a *wh*-element and C has both uT and uWh, the subject movement can check both uT and uWh while T-to-C movement checks only uT, requiring another instance of *wh*-movement. In this particular circumstance, the subject movement is more economical.

**Problems:** In spite of its elegance, P&T’s analysis has some problematic aspects. First, their account requires positing that complementizers are T elements. Thus, any appearance of a complementizer *that* should be a consequence of T-to-C movement. Second, they argue that the absence of *that* in (3) involves subject movement to Spec, CP.

(3) I think [CP Leslie C [TP <Leslie> ran for public office]]. However, it is hard to find independent evidence showing that the embedded subject Leslie resides in [Spec, CP] in (3) but it stays in [Spec, TP] if a complementizer *that* is present. Third, if both *that* and *do*-support are the result of T-to-C movement, their account cannot explain the negative inversion case in (4), in which *that* and *do*-support co-exist.

(4) I think that only in that election did Leslie run for public office. (Rizzi 1997)

The purpose of this paper is to improve on P&T while maintaining an account that simultaneously captures both the *that*-t effect and the *do*-t effect, which have been infamous puzzles in syntactic theory.

**Analysis:** A simple intuition is pursued in this paper. Namely, the reason why both moving T to C and merging *that* cause ungrammaticality is because there is no proper slot for MOVE or MERGE in terms of phrase structure. I assume that the position for *that*-insertion and T-to-C movement target the Fin in cartographic CP structures (Rizzi 1997). The generalization is that both T-to-C movement and *that*-insertion cannot occur when there is a *wh*-phrase in [Spec, TP]. It can be schematized as in (5).

(5) *[FinP Fin [TP wh T’]]

Under (5), the analysis of (1) and (2) follows trivially. Both the embedded TP of (1) and the matrix TP of (2) have a *wh*-phrase in [Spec, TP]. Consequently, what is forced in terms of phrase structure building is the absence of Fin, the proper slot for *that* in (1) or *did* in (2). The appearance of *that* or *did* causes ungrammaticality of (1) or (2) due to the lack of the slot. Since this analysis does not depend on Case, it can be readily extended to the *that*-t effect in a locative inversion (LI) sentence in (6) and the *for*-t effect in (7).

(6) [Into which room] do you think (*that) t entered Bill? (the *that*-t effect in LI)
(7) [Which team] do you prefer (*for) t to win? (the *for*-t effect)

**The Absence of Fin:** One question immediately arises: how can we derive the absence of Fin?

To answer this question, two ingredients are implemented: functional hierarchy and labeling.

(8) Force > (Top) > (Foc) > (Top) > Fin
First, I assume that Rizzi’s CP hierarchy in (8) shows selection properties among functional projections. Given that Top projection is optional, Foc can c-select FinP but not vice versa. Once TP is labeled as a higher projection than Fin due to the wh-phrase in [Spec, TP], Fin cannot participate in the derivation.

Second, this paper pursues a hybrid approach on labeling; we adopt Cecchetto and Donati (2015) regarding the definition of labeling, as in (9a), and Chomsky (2013) regarding the mechanism of labeling, as in (9b).

(9) a. A label can trigger further computation and is visible from outside the syntactic object. (Cecchetto and Donati 2015)

b. If there is an SO = {XP, YP}, the identical feature in both XP and YP is the label of the SO. (Chomsky 2013)

What (9a) tells us is that labeling plays an important role in narrow syntax; labeling involves selection in the syntactic derivation. (9b) tells us that the label of a syntactic object is intersection of two constituents. For the current purpose, TP is labeled as FocP iff (i) wh-phrases involve [Foc]-feature (Rizzi 2004) and (ii) T involves [Foc]-feature (Bošković 2002) so that T, as a head, has become the label of T'. (10) shows the comparison between standard labeling (cf. (10a)) and the labeling proposed in this paper (cf. (10b)). Under the functional hierarchy in (8), (10b) cannot merge with Fin since Foc is higher than Fin in the hierarchy.

(10) a. [TP wh [T vP]]

b. [FocP wh[Foc] [T[Foc] T vP]]

Note that the labeling TP as FocP, as in (10b), appears only when there is a wh-phrase in [Spec, TP] given that non-wh-subject does not involve [Foc]-feature, which explains the famous asymmetry between subject and object/adjunct with respect to the that-t effect.

Variation: As Sobin (1987) notes, the that-t effect exhibits dialectal variation. To explain this, Boeckx (2008) and Lohndal (2009) claim that English complementizer that can be lexicalized either as Force or as Fin. The speakers who allow (1) with that lexicalize that as Force. According to Lohndal (2009), the Force-complementizer account also capture Polish, which does not have the Comp-t effect. Their account can be adopted to the current analysis without any cost. Although Fin is absent in the phrase structure, that can still be merged in Force. The fact that that is potentially merged in various positions might be desirable since it can explain (4); T-to-C movement targets Fin whereas that is merged in higher position than Fin. This analysis can also account for the difference between the that-t effect and the do-t effect; unlike the that-t effect, the do-t effect does not have dialectal variation.

Conclusion: As an alternative to P&T, this paper suggests that the that-t effect and the do-t effect can be captured by the absence of Fin with respect to phrase structure building.

A Model-Theoretic Approach to Minimalist Syntax

Yiding Hao

Abstract

Modern Minimalist theories of syntax view grammar as a *procedure* in which words are combined to form phrases in a systematic way. This is done using *operations* such as merge and move, which produce complex expressions based on properties of the inputs known as *features*. On the other hand, the Principles and Parameters framework that underlies Government and Binding Theory views grammar as a set of descriptive statements that restrict the set of possible syntactic structures in a language. Under this view, a sentence is grammatical if and only if it is accurately described by all the constraints.

These two approaches to syntax—thinking of grammar as a procedure and thinking of grammar as constraints on permissible structures—are known as *derivational* and *constraint-based* syntax, respectively. At first glance, derivational and constraint-based syntax seem to be radically different views on how syntactic analysis can be conducted. However, in this paper I argue that for Minimalism, these two approaches are formally equivalent. In particular, I will show that for any (derivational) minimalist grammar, there is a constraint-based grammar generating the same language.

This will be done by formalizing minimalist grammars as sentences in monadic second-order (MSO) logic, an extension of first-order logic that allows quantification over unary predicates. The seminal work of [Rogers, 1998], which stated the constraints of Relativized Minimality in MSO logic, showed that MSO logic can be seen as a formalization of constraint-based syntax. Furthermore, work in mathematical logic has shown that there is a close connection between MSO logic and derivational approaches to syntax. In particular, a set of trees is definable in MSO logic if and only if it is recognizable by a deterministic bottom-up tree automaton, a model of computation that can be used to formalize derivational frameworks.

Derivational Minimalist syntax will be formalized using *minimalist grammars*, as defined in [Stabler, 1997]. In this formalism, the merge and move operations apply deterministically whenever the appropriate feature-checking configuration is achieved. Minimalist grammars belong to the class of *mildly context-sensitive* grammars, a class of formalisms that can generate the context-free languages as well as languages with cross-serial dependencies such as the variety of Swiss German described in [Shieber, 1987]. Minimalist grammars have been shown to be equivalent in generative capacity to multiple context-free grammars and linear context-free rewriting systems.

The MSO characterization of minimalist grammars will follow the *Two-Step Approach* of [Morawietz, 2008]. In general, constituency trees produced by a mildly context-sensitive formalism are not MSO-definable. Instead, the Two-Step Approach decomposes the formalism into two parts: the *derivation tree*, a structure encoding the sequence in which operations are applied to construct a sentence, and a *transduction* that maps the derivation tree to the derived constituency tree by carrying out the operations encoded in the derivation tree. For minimalist grammars, both the derivation trees and the transduction from derivation trees to constituency trees are MSO-definable.

The MSO sentence defining the minimalist derivation trees is constructed from [Kobele et al., 2007], which constructs a deterministic bottom-up tree automaton that recognizes derivation trees. Since tree automata are equivalent to MSO logic, this automaton is converted into a sentence in MSO logic. The transduction from derivation trees to constituency trees is constructed by observing that the derivation trees are isomorphic to the constituency trees before movement has been applied.

The construction given in this talk shows that every minimalist grammar can be represented as a sentence $\varphi$ in MSO
logic. In addition, [Kobele, 2011] and [Graf, 2011] show that for any MSO sentence $\psi$, the sentence $\varphi \land \psi$ is also a representation of a minimalist grammar, up to relabelling the interior nodes. This means that minimalist grammars and MSO constraints are interchangeable: a minimalist grammar can be combined with MSO constraints to produce a new minimalist grammar, and MSO constraints can be combined with minimalist grammars to produce new MSO constraints.

To conclude, we outline a new methodology for syntactic investigation in a Minimalist framework arising from these results. When encountering new data, the researcher must try to give an analysis of the phenomena under consideration using an existing grammar $\varphi$. If this is impossible, the researcher can stipulate changes $\psi$ in either the minimalist grammar or the constraints—whichever is easier. The original grammar and the new stipulations are combined to form a new grammar $\varphi \land \psi$. This new grammar can either be viewed as a constraint, or it can be converted back into a minimalist grammar. In this way, derivational minimalist grammars can be freely combined with P&P-style constraints to produce new syntactic analyses.

References


Oblique causer constructions in Lezgian
Philip Shushurin

Data. Many languages of the world possess constructions where an external oblique argument is interpreted as an agent with no volitional control over the situation – the oblique causer construction (OblCC). See the following German example:

(1) Die Vase zerbrach mir.
The vase broke. I.DAT
‘I accidently broke the vase’

According to Schäfer (2008), the OblC in such constructions are compatible with three interpretations: ‘accident’ reading (see e.g. (1)), ‘failed prevention’ reading, ‘success with difficulty’ reading. Additionally, such constructions are subject to the following constraints: they require change-of-state predicates and they an animate oblique causer. Lezgian, a Dagestani language spoken in the Caucasus, possesses a such construction, see (2):

K-ADEL A.(ABS) home-DAT bring-NMNLZ BECOME-AOR
‘Kerim unintentionally brought Ali (home)’

In this construction the unintentional causer is marked adelative case (‘from X’), the use of intransitive auxiliary χana is obligatory, the verb (which can transitive or intransitive) appears in the nominalized form. The OblC are compatible with three readings cited in Schäfer (2008). However, such constructions in Lezgian are different from languages discussed in Schäfer (2008) in at least two important respects. First, in Lezgian, unlike German, Icelandic (Wood 2013) and oth., OblCCs are incompatible with transitive predicates, as long they are change-of-state. Secondly, OblC in such constructions has subject properties, thus, it can control PRO and reflexives (unlike German).

Analysis. I want to propose that in Lezgian the intransitive auxiliary χana can take either a DP or a nominalized clause as its internal argument. The nominalized clause consists of a nP taking a voiceP as its complement. Here is the derivation of (2):
1. The transitive clause which contains a voice is embedded under a nP which creates a nominalized clause;
2. This nominalization is merged as an internal argument of the intransitive auxiliary *hana;
3. In the matrix clause the voice phrase is non-active, i.e. it enters the derivation with valued phi-features (perhaps default), its head is occupied by the semantically vacuous change-of-state auxiliary *hana; its specifier is not merged;
4. An applicative head is merged with voiceP and in its specifier a PP is merged which contains an OblC;
5. The PP containing the OblC moves to SpecTP;

I take an agnostic position on whether the voiceP contains a PRO in its specifier, or is devoid of subject whatsoever. In the first case I assume the PRO is controlled by the adative argument of the matrix clause; in the second case I would assume that the two clauses are connected via voice incorporation (see Wurmbbrand 2014). Furthermore, I want to propose that the matrix voice phrase does not contain an argument, rather, it merges an applicative projection, which introduces a prepositional phrase containing the unintentional causer. This argument gets its theta (CAUSER) from the head of the prepositional phrase (I assume that adative-marked nominals are PPs). One may ask why UC constructions are incompatible with normal active morphology, in other words, why the Lezgian versions of sentences of the following type are disallowed:

(3) * from.me hit John  
int. meaning ‘John hit and I caused it unintentionally’
(4) * from.me Bill hit John  
int. meaning ‘Bill hit John and I caused this happen unintentionally’

I assume, following Legate (2014), that an active voiceP enters the derivation with unvalued phi-features, and for the derivation to go further, the voice projection must value its phi-features locally, which can be fulfilled by merging a specifier (the subject). The ungrammaticality of 3 is thus due to the fact that the voiceP has unvalued phi-features. The ungrammaticality of (4) is, I want to argue, of semantic nature: once an argument has a agent theta-role, no other argument can bear the causer theta-role. (I leave open the question whether theta roles ‘causer’ and ‘agent’ are independent or can be derived through one another). I also assume that the requirement for the verb to be change-of-state is also of semantic nature: the theta-role CAUSER is incompatible with non-change-of-state predicates.

**Deriving the contrast between Lezgian and German.**

To recall, there are two major differences: the subject properties in Lezgian vs. the lack thereof in German and the constraint on transitive predicates in German vs. the lack thereof in Lezgian. I assume two main differences between Lezgian and German: (1) in Lezgian and not in German transitive clauses can be embedded under nominalizations; (2) in Lezgian and not in German PPs can percolate their phi-features up. I also assume that in order for the derivation to go through, a TP of the clause must valuate its phi-features. In German the only way to do so is to move the internal argument of the intransitive verb (in example (1): die Vase) to the SpecTP. In Lezgian another option is available. The external argument of the intransitive auxiliary can be merged as a nominalization. The TP, which needs its phi-features, gets them from the OblC which moves to SpecTP. It explains both the availability of transitive predicates in OblCCs and the subject properties of OblCs.

**Bibliography:**

Mandarin Chinese exhibits a restriction on the co-occurrence of certain aspect- and modality-related adverbs and sentential clitics (with the latter often referred to in the Chinese literature as sentence final particles). As seen in (1), the adverbs yijing ‘already’ and jiu ‘just.then’ co-occur with le and disallow ne, while the adverbs hai ‘still’ and cai ‘only.then’ co-occur with ne and disallow le.

(1)  
   a. Xianzai yijing wudian le (*ne)  
      now already 5.o’clock LE  
      ‘It’s already five o’clock’  
   b. Wo mingtian jiu qu le (*ne).  
      I tomorrow just.then go LE.  
      'I'm leaving tomorrow (so there's no way I can accept your invitation)'  
   c. Ta hai xiao ne (*le).  
      He still small NE  
      He’s still young (e.g., so you can’t expect him to know that word)  
   d. Wo mintian cai qu ne (*le)  
      I tomorrow only.then go NE  
      'I'm not leaving till tomorrow'

The complementary distribution of Mandarin le and ne suggests that the two are variants spellouts of the same functional head. I take this head to be located within the low end of the expanded CP domain, given its sensitivity to material lower in the clause, and follow Li (2006) in calling it Deik° (for deixis). Applying the framework of Pesetsky and Torrego (2007), I assume Deik° is merged with the uninterpretable, unvalued feature [uDeik:__]. When c-commanded by yijing ‘already’, which contains the interpretable feature [iDeik:proximal], Deik° gets valued as [Deik:proximal]. In this case, it spells out as le and encodes aspectual anteriority and a temporal (or potentially discourse-pragmatic) sense of closer-than-you-think. Alternatively, when valued as [Deik:distal] by the adverb hai ‘still’, it spells out as ne and encodes continuative aspect and a farther-than-you-think meaning.

Antisymmetry requires that the underlying structure of the sentences in (1) be [DeikP [TP]]. Subsequently, in order for Deik to be valued by the relevant adverb, the entire TP raises into Spec,DeikP - a case of clause-level pied-piping. No doubt, a more direct means of achieving the same feature valuation would be to simply raise the adverb itself. That this is not attested can be attributed to a general prohibition in Chinese against phrasal movement – the same principle that requires that wh-phrases in Chinese remain in-situ in questions, rather than raising to C. Covert movement of the adverb into Spec,DeikP would have likewise obviated the need for TP-pied-piping. But the curious well-formedness of a sentence like (3), despite the fact that it involves a wh-word within a complex NP island, suggests that Chinese wh-words achieve their sentential scope via operator binding, rather than covert movement to C.

(3) Akiu kanbuqi [zuo shenme de ren]? [Tsai 1999:42]  
      Akiu despises does what REL person  
      ‘*What does Akiu despise a person who does (for a living)’
With overt and covert raising of the adverb both ruled out, the only available choice is pied-piping of the containing TP, not unlike the way English pied-pipes the entire DP in (4) in order to bring the [wh] feature up to C.

(4) [[Whose father’s] book] did you buy?

One prediction made by this hypothesis is that there should be other instances of adverb-C dependency in Chinese apart from DeikP. Pairs such as hui ‘certainly’ + de (Fin°) and yinggai ‘probably’ + ba (Force°) bear this out. Another prediction is that in languages without restrictions on phrasal movement, sentential clitics should surface in a more transparently left-peripheral position. The sentences in (5) show that this holds for Tagalog. Here, Deik is spelled out as na when valued as [proximal] and as pa when valued as [distal], and consistently occurs in second-position (with the element in first position being either a head-raised verb or topicalized phrase).

(5) a. Sa Biernes na ang piesta.
   on Sunday NA the party
   'The party is next Friday (and there is little time between now and then)

   b. Sa Biernes pa ang piesta.
   on Sunday PA the party
   'The party is next Friday (and there is lots of time between now and then)

Unlike Chinese, Tagalog freely permits phrasal- and head-movement. It also lacks overt deictic adverbs. Instead, a proADV is generated low in the clause and raises to the specifier of Deik, valuing it accordingly. (Thus, sentential clitics can be understood along the same lines as Sportiche’s (1996) analysis of pronominal clitics as functional heads agreeing locally with null pronouns.) Further support for this kind of adverb-C agreement as a general syntactic mechanism is found in Czech, where the so-called “fringe clitic” už ‘already’ marks proximal deixis, and can optionally spell out either an adverb low in the clause, or the Deik-head high in the clause into whose specifier that adverb’s null alter ego has raised for feature valuation.

(6) A ten mi oznámil, že (už) jsem si tě (už) najal.
   And he me informed, that (deik:proximal) have.1sg ref.DAT you.ACC (already) hired
   ‘And he informed me, that I already hired you.’

References
In Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation. Amsterdam: Benjamins.

1 This is not to say that adverbs such as yijing 'already' and hai ‘still’ cannot have covert counterparts. On the contrary, it is likely they do, given that their overt presence is not needed to license le or ne.
Homeostatic Reinforcement Learning for Harmonic Grammars
Jon Rawski

Background. Reinforcement Learning (RL), despite being one of the most widely used and neurologically robust learning algorithms for finite-state automata, has an uneasy history with linguistics. Specifically, the requirement of an internal, restricted hypothesis space and other learnability restraints is inadequately satisfied by externally defined "naive" reward (Chomsky 1959). Parallel difficulties exist in computational neuroscience, where the “reward” required by RL algorithms is often defined ad-hoc or extremely oversimplified.

Reparation of RL and linguistics is made even more urgent by the discovery that: 1) phonology is at most a regular language (Heinz 2011), meaning it is restricted to finite-state automata, and 2) RL is perfectly computed by cortical neurons (Schultz et al 1999). One recent attempt is Charles Yang’s (2002) “Naïve Parameter Learner”, which uses RL to successfully model acquisition of overt [WH-movement] and [V2] parameters, yet fails to provide more than an ad-hoc definition for “reward”. Clearly, a language learning model requires a reward definition which is derived from the internal grammar’s characteristics, as well as a dynamic link between the internal state and the external actions.

Homeostatic Reinforcement Learning

In this work I show that recent insights from computational neuroscience offer a possible strategy. A recent framework called Homeostatic Reinforcement Learning (HRL) (Keramati and Gutkin 2014) treats “reward” as an internal satisfaction of multiple, parallel constraints in a homeostatic space (Fig 1). Homeostasis is the tendency of biological systems to maintain their internal structure. In HRL, deviations from an internal setpoint in this space decrease the rewarding value of an action, and thus define the acceptable paths between states in a finite state automaton. Such a framework also has the benefit of being neurally attested.

Satisfaction and violation of parallel constraints immediately brings to mind Harmonic Grammar (Smolensky and Legendre 2006), where grammars are defined via candidates interacting with a set of internal weighted, parallel constraints. In Harmonic Grammar (HG), learning occurs by changing the constraint weights. Here I posit that the weighted constraints in Harmonic Grammar constitute a homeostatic space, and that the Harmony function generates a local minimum that is a necessary and sufficient condition for RL in constraint-based grammar learning.

Learning Experiment: Final Consonant Devoicing

Whether a language restricts voiced obstruents in syllable codas via devoicing or not can be captured by two violable constraints:

* Coda-Voice: Every obstruent in coda position is voiceless
Ident-Voice: Every consonant's voicing specification is identical in input and output

The deletion depends on the constraint weightings, where \( w(*\text{Ident-Voice}) > w(*\text{Coda-Voice}) \), or vice versa. This behavior is seen in languages like Russian, where the underlying consonant devicles in the Nominative form, but not in Genitive, where the added suffix \(-a/\) protects the underlying consonant from devoicing

(1) Russian final devoicing

<table>
<thead>
<tr>
<th>Nominative</th>
<th>Genitive</th>
<th>English gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>/glaz/</td>
<td>/glaz-a/</td>
<td>‘eye’</td>
</tr>
<tr>
<td>[glas]</td>
<td>[glaza]</td>
<td></td>
</tr>
</tbody>
</table>

The former grammar will generate a Harmony tableau for the input /bad/ as in Figure 2a. Learning this grammar in a Harmonic Grammar framework requires adjusting the weights based on input data. I will show that, starting from any ranking, a learner using the HRL framework correctly converges on the target values to maintain an internal grammar state (fig 2b).

![Figure 2](image)

**Figure 2.** (a) Constraint tableau for a coda obstruent devoicing grammar. (b) learning simulation of (a) in an HRL framework.

**Conclusion.** Though Harmonic Grammar was the first linguistic theory to attempt to derive linguistic structure from neural principles, it has not been so principled in deriving language learning. In this work I argue that the HRL approach offers prospects for unifying ideas from neural and linguistic theory in order to provide a more coherent explanatory neurolinguistics.

**References.**
TextPros: Prosody-text aligner for unlabeled speech

Hussein Ghaly

The goal of this research is to develop a tool (TextPros) for automatic detection of prosodic boundaries within unlabeled speech material and aligning them with its text. Most automatic systems are trained statistically through Machine Learning techniques using labeled data, to be able to detect the prosodic boundaries in novel test data (Rosenberg, 2009; Garrido, 2013; Mertens and Simon, 2013). Few other approaches use unlabeled audio files as their input, such as Ludusan and Dupoux (2014). To measure their accuracy, all of these approaches were tested on labeled audio material, such as Boston University Radio News Corpus (BURNC), or audio files annotated by trained phoneticians. Results have shown that methods using labeled material are more accurate (Ludusan and Dupoux, 2014). However, it is still important to develop automatic approaches for detecting prosodic boundaries without labeled material, which is not readily available, especially for low-resource languages.

The input material for our tool is unlabeled audio files with their corresponding text. Initially, the sentences examined with this tool are double center embedded sentences in French. The prosody of this type of sentences was previously studied in English (Fodor and Nickels, 2011) and in French (Desroses, 2014). These sentences are in the form of NP1 NP2 NP3 VP1 VP2 VP3, so the goal is to automatically detect the prosodic boundaries at the five locations between these phrases. The audio files examined are 545 files recorded by 30 speakers for 36 sentences, and the test material consists of a subset of the audio files (48 files), with the prosodic boundaries marked on the sentence text by a judge who is a native speaker of French.

Since there is no annotation in the audio files for word/phrase boundaries, neither for the tones and break indexes (ToBI), it is necessary to automatically align the words/ phrases from the sentence text to time intervals of the audio file. This was achieved first by automatically identifying characters with voiceless sounds in the IPA string of the sentence text (which is obtained using an online converter from French to IPA), and matching them with discontinuities of pitch/F0 detected by Praat software, leading to a set of potential alignment points between text and audio.

In addition, the five locations between the syntactic phrases in the sentence text are potential locations for prosodic boundaries in the audio file. The cues for prosodic boundaries considered here are silent pauses of duration greater than 250 milliseconds and sharp pitch falls, (falling at a rate of 10 Hz per 10 milliseconds). These pitch falls may correspond to the transition from H* to aL in F_ToBI at the start of Accentual Phrase (AP) (Delais-Roussarie et al., 2015). The mapping between prosodic cues and locations of the syntactic boundaries within text provides additional alignment points. The algorithm uses dynamic programming to get the most likely (linear) path between these points, hence predicting whether there is a prosodic boundary at each location.
Applying the tool to the test material, accuracy was 47.6%, indicating to what extent it is able to detect whether or not there are prosodic boundaries at the locations indicated by the human judge, while the recall was 61.2%. There was resemblance between the overall distribution of prosodic boundaries detected automatically and by the human judge, as shown in figure 1. The discrepancy can be due to the overlapping in French between voiceless /k/ in the complementizer “que” and the syntactic phrase boundaries at locations 1 and 2, so the algorithm tends to ignore boundaries there and register the boundary at location 3 instead. Better future results can be obtained by optimizing the alignment algorithm, better accommodating prosodic style variation for individual speakers, and using the agreement points of multiple human judges. Different languages and sentence structures can be examined with this tool in the future.

Figure 1 - Prosodic boundaries automatically detected at the five locations of syntactic phrase boundaries vs. those indicated by human judges, for two kinds of sentences with different patterns of their phrase lengths (A & B, assumed +/- prosodically helpful respectively)

References
Tone analysis in Tai Khamti: computational models in language documentation

Rikker Dockum

Study of tone systems is fundamental to language documentation in Southeast Asia and many areas of the world. There is an ongoing need for better training and better documentation methods, as evidenced and addressed in collections like Bird and Hyman (2014). Meanwhile, computational methods are opening up possibilities for studying tone systems in novel ways. A quantitative study by Brunelle and Kirby (2015) confirms that the simplistic tonal-atonal dichotomy is a poor representation of the complexity of tonal typology in Southeast Asia, and also that vertical transmission is crucial, as language family was the strongest predictor of the typological profile of a given language’s tone system. Improving documentation of tone systems thus has clear benefits to for both synchronic and diachronic linguistics. Shosted et al (2014, 2015) in their work with Iu Mien [ISO 639-3: ium] argue for implementing unsupervised computational modeling of lexical tone very early in the documentation process for understudied languages. Results such as these motivate the expanded use of computational methods in conjunction with traditional language documentation methods.

This study implements and expands on methods outlined in Shosted et al (2014) for a previously undocumented variety of Tai Khamti [ISO 639-3: kht] spoken in Khamti Township, Sagaing Region, Burma. Tones on target words were manually segmented and extracted into two test corpora of citation forms and running speech. The time- and pitch-normalized data were then modeled using principal components analysis (Johnson 2008:95) and k-means clustering in R (R Core Team 2015). Tones were modeled using both exclusively pitch data and pitch data plus various phonation measures (spectral tilt, jitter, shimmer). These models were then compared against the author’s earlier work assigning tone categories using traditional methods of researcher audition and instrumental analysis. Current results indicate that using only pitch data from citation tones, this unsupervised clustering agreed with the categories identified through traditional methods with very high precision, between 0.93 and 1.0, in three of the four expected tonal categories, and recall 0.79-0.86 in all four categories. Additionally, the introduction of spectral tilt measures improved recall on the low tone from 0.79 to 0.86, which indicates non-modal phonation as a potential cue for this tone, adding empirical weight to the existing impressionistic judgments of the author.

Such comparison enables mutual methodological improvement. Computational modeling can extend and strengthen existing tonal analyses, where fieldworker expertise often varies. Moreover, linguists experienced in traditional methods can help to further delineate and refine the usefulness of the quantitative methods, and explore new areas of linguistics to apply them in. Refinement of these methods in turn achieves a more sophisticated and increasingly nuanced understanding of tonal phenomena in Southeast Asia and beyond.

References


LOCAL SERVICES: FOOD & PRINTING

**Quick lunch**
Oxford Café
375 Fifth Ave. (between 35th and 36th St.)
Sandwiches, salads, pastas, soups - $

Chipotle
34th St. (between 5th and 6th Ave.)
Burritos, tacos - $

**Indian**
Minar
5 W. 31st St. (between 5th and 6th Ave.)
Indian buffet - $

**Korean**
Woorijip
12 W. 32nd St. (between 5th and 6th Ave.)
Cafeteria-style Korean restaurant - $

Seoul Garden
34 W. 32nd St. (between 5th and 6th Ave.)
Vegetarian Korean food - $$

Mandoo Bar
2 W. 32nd St. (between 5th and 6th Ave.)
Korean dumpling restaurant - $$

**Japanese**
Café Zaiya
18 E. 41st St. (between 5th and Madison Ave.)
Japanese Bento boxes - $

Sushi Sennin
30 E. 33rd St. (between Madison and Park Ave.)
Sushi - $$$

**Latin American**
Leña Latin Grill
34 E. 35th St. (between 5th and 6th Ave.)
Latin-style wraps and burritos - $$

Tina’s Cuban Cuisine
179 Madison Ave. (between 33rd and 34th St.)
Cuban sandwiches and plates - $

**Middle Eastern**
Nanoosh
173 Madison Ave. (between 33rd and 34th St.)
Mediterranean lunch options - $

**Copy shops/printing services**
FedEx
350 Fifth Ave. (in the Empire State Building)

Staples
16 E. 34th St. (between Madison and Fifth Ave.)
MAP OF THE AREA

Google Maps. (2012). [SYNC No. 14, 2012, New York, NY] [Street map]. Retrieved from https://maps.google.com/maps/ms?msid=210384596724109246865.0004c9d92961e06c97c7e&msa=0&ll=40.74896,-73.984192&spn=0.002434,0.003482
## SYNC 2016 PARTICIPANTS

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christina Tortora</td>
<td>CUNY Graduate Center</td>
<td><a href="mailto:CTortora@gc.cuny.edu">CTortora@gc.cuny.edu</a></td>
</tr>
<tr>
<td>Wafa Bahri</td>
<td>CUNY Graduate Center</td>
<td><a href="mailto:wafafedy31@gmail.com">wafafedy31@gmail.com</a></td>
</tr>
<tr>
<td>Taletha Callahan-Kanik</td>
<td>CUNY Graduate Center</td>
<td><a href="mailto:tally.callahan@gmail.com">tally.callahan@gmail.com</a></td>
</tr>
<tr>
<td>Rikker Dockum</td>
<td>Yale University</td>
<td><a href="mailto:rikker.dockum@yale.edu">rikker.dockum@yale.edu</a></td>
</tr>
<tr>
<td>Hussein Ghaly</td>
<td>CUNY Graduate Center</td>
<td><a href="mailto:hmghaly@gmail.com">hmghaly@gmail.com</a></td>
</tr>
<tr>
<td>Yiding Hao</td>
<td>New York University</td>
<td><a href="mailto:yiding.hao@yale.edu">yiding.hao@yale.edu</a></td>
</tr>
<tr>
<td>Hwichan Oh</td>
<td>Stony Brook University</td>
<td><a href="mailto:hwihan.oh@stonybrook.edu">hwihan.oh@stonybrook.edu</a></td>
</tr>
<tr>
<td>Jon Rawski</td>
<td>Stony Brook University</td>
<td><a href="mailto:rawski.jon@gmail.com">rawski.jon@gmail.com</a></td>
</tr>
<tr>
<td>Mary Robinson</td>
<td>NYU</td>
<td><a href="mailto:mkr361@nyu.edu">mkr361@nyu.edu</a></td>
</tr>
<tr>
<td>Philip Shushurin</td>
<td>NYU</td>
<td><a href="mailto:phshushu@gmail.com">phshushu@gmail.com</a></td>
</tr>
<tr>
<td>Russell Tanenbaum</td>
<td>Stony Brook University</td>
<td><a href="mailto:russell.tanenbaum@stonybrook.edu">russell.tanenbaum@stonybrook.edu</a></td>
</tr>
<tr>
<td>Ji Yea Kim</td>
<td>Stony Brook University</td>
<td><a href="mailto:jiyea.kim@stonybrook.edu">jiyea.kim@stonybrook.edu</a></td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

The SYNC 2016 Organizing Committee:
Matthew Stuck (co-chair), Ivana Đurović (co-chair), Adina Williams (NYU), Dan Duncan (NYU), Luke Lindemann (Yale), Paola Cepeda (Stony Brook)

Special thanks to:
Gita Martohardjono, Nishi Bissoondial, Christina Tortora, Ian Phillips

SYNC 2016 abstract reviewers:
Suzy Ahn, Honaihad Ahyad, Hagen Blix, Sara Catlin, Paola Cepeda, Aniello de Santo, Dan Duncan, Masha Esipova, Viviana Haase, Paloma Jeretic, Ivana LaTerza, Christen N Madsen II, James Monette, Yining Nie, Ian Phillips, Vahideh Rasekhi, Nazila Shafiei, Allison Shapp, Dunja Veselinovic, Adina Williams, Hongchen Wu, Linmin Zhang, Muye Zhang

Participating programs:
M.A./Ph.D. Program in Linguistics
The Graduate Center, City University of New York

Department of Linguistics
New York University

Department of Linguistics
Stony Brook University

Department of Linguistics
Yale University

And all of our volunteers and session chairs!