PROCEEDINGS OF THE FOURTY-FOURTH ANNUAL MEETING OF THE

BERKELEY LINGUISTICS SOCIETY

February 9-11, 2018

Editors

Karee Garvin

Noah Hermalin

Myriam Lapierre

Yevgeniy Melguy

Tessa Scott

Eric Wilbanks

Berkeley Linguistics Society
Berkeley, CA, USA

Berkeley Linguistics Society
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720-2650
USA

ISSN: 2377-1666

LCCN: 76-640143

Contents

Acknowledgements .... **iii**

Foreword .... **v**

I can believe it: Quantitative evidence for closed-class category knowledge in an English-speaking 20- to 24-month-old child
Alandi Bates, Lisa Pearl, & Susan R. Braunwald 1

Quantitative Comparison for Generative Theories: Embedding Competence Linguistic Theories in Cognitive Architectures and Bayesian Models
Adrian Brasoveanu & Jakub Dotlačil 17

On the Paradox of Changting Hakka Tone Sandhi
Luhua Chao 33

The pragmaticalization and synchronic variations of Japanese adverb jitsuwa and English adverb actually
Kiyono Fujinaga 47

Subanon mo- cancels out volitionality: Evidence from paradigms and argument structure
Bryn Hauk 61

Two types of preverbal object movement and duration/frequency phrases in Mandarin Chinese
Hsi-Lun Huang 83

Effects of sensory modality on the interpretation of subjective adjectives: Comparing sight, smell and taste
Elsi Kaiser 99

Gradable Ideophones, Scales, and Maximality in Grammar
Koji Kawahara 115

Suppletive allomorphy conditioned by humbleness in Korean
Soo-Hwan Lee & Minjung Kim 131

A feature-based analysis of the Chol (Mayan) person paradigm
Carol-Rose Little 147

The attributivizing marker in Andi (Nakh-Daghestanian): clitic, affix, or both?
Timur Maisak 163

Response Bias and Sensitivity in Distributional Learning: Two Stages in Early Phonetic Category Acquisition
Emily Moeng 179

Interpretation of internally-headed relative clauses in Shan
Mary Moroney 197

An anaphora-based review of the grammar/pragmatics division of labor
Koichi Nishida 213

Underspecification and scope: The case of gapping
Sang-Hee Park 229

Phonological environment and the social evaluation of American English sibilants
Jacob Phillips 243

Syllabic Size Restrictions on Verb Reduplication in Brazilian Portuguese
Jamilläh Rodriguez 257

Viewpoint aspect and object case in $K^w ak^w ala$ and Finnish
Katie Sardinha 271

Japanese -TE as a marker of aestheticization in the language of menus
Kiyoko Toratani 287

Parameter dependencies: Towards descriptive adequacy for phonological accent systems
Alexandre Vaxman 301

Acknowledgements

The Executive Committee of the 44th Annual Meeting of the Berkeley Linguistics Society would like to thank the students, faculty, and staff who gave their time to review abstracts, chair sessions, and assist with a vast number of administrative and logistical tasks–this conference would not have been possible without you. We also extend our gratitude to conference participants, whose contributions have once again made this conference an engaging and stimulating experience. Special thanks go to Paula Floro and Belén Flores, whose knowledge, expertise, and guidance proved indispensable in making BLS44 a success. Finally, we thank the UC Berkeley Linguistics Department for its financial assistance.

Foreword

This collection contains 20 of the talks given at the 44th Annual Meeting of the Berkeley Linguistics Society, held in Berkeley, California, February 9-11, 2018. The conference focused on the special topic of innovative methods, and featured presentations on a broad range of theoretical, experimental, and descriptive work, especially with regards to innovative methodologies.

BLS44 was organized by the second-year graduate students of the Department of Linguistics at the University of California, Berkeley: Karee Garvin, Noah Hermalin, Myriam Lapierre, Yevgeniy Melguy, Tessa Scott, and Eric Wilbanks. The papers in this volume were edited for style by the members of the Executive Committee, and compiled into the final volume by Noah Hermalin and Yevgeniy Melguy.

The BLS 44 Executive Committee
January 2019

I *can* believe it:
Quantitative evidence for closed-class category knowledge
in an English-speaking 20- to 24-month-old child

ALANDI BATES, LISA PEARL, & SUSAN R. BRAUNWALD
University of California, Irvine[*]

1 Introduction

Adults are believed to have abstract syntactic categories that they use to generate their observable utterances (e.g., closed-class categories like NEGation and open-class categories like VERB to generate *don't go*). However, there's been significant debate about when children develop syntactic categories and how to accurately assess what category knowledge they have when. We review prior approaches to assessing children's developing knowledge of syntactic categories, and then present our quantitative approach, which synthesizes insights from this prior work. This allows us to (i) define possible child representations for multi-word combinations, and (ii) calculate the observed vs. expected linguistic production properties for each possible representation. We use this approach to investigate the existence of both closed-class and open-class syntactic categories in a 20- to 24-month-old child's verb phrases. We evaluate whether the child's observed production matches the expected production when the child uses a specific category representation, and find that the child's productions are compatible only with representations that have adult-like closed-class categories (NEG, AUXiliary), but not adult-like open-class categories (NOUN, VERB). We conclude with implications for the development of syntactic categories.

2 Syntactic category knowledge in children

There hasn't been a clear consensus for when children develop syntactic categories, whether open-class or closed-class. Some studies suggest that knowledge of certain categories – either rudimentary or adult-like – may be in place as early as age 2 Pinker (1984), Valian

[*]We have benefited from comments and suggestions from the Computation of Language Laboratory at UC Irvine, and the audiences at UROP 2017 at UC Irvine, CAMP 2017 at UCLA, and the Berkeley Linguistics Society 2018 meeting. All errors are of course our own and not at all their fault.

(1986), Capdevila i Batet & Llinàs i Grau (1995), Booth & Waxman (2003), Rowland & Theakston (2009), Theakston & Rowland (2009), Yang (2010, 2011), Shin (2012), Meylan et al. (2017), while others argue that such knowledge only emerges much later (Pine & Lieven 1997, Tomasello 2004, Kemp et al. 2005, Tomasello & Brandt 2009, Theakston et al. 2015). Taken together, there seems to be some agreement that children may have rudimentary knowledge of open-class categories (NOUN, ADJective) fairly early, but don't refine these into adult-like open-class categories until later. However, for closed-class categories (DETerminer, NEG, AUX), there isn't yet consensus on when either rudimentary or adult-like versions of these categories develop. This may be due in part to the different quantitative analysis approaches that prior research has adopted. More generally, many prior studies demonstrate that there's utility in quantitatively analyzing children's productions to determine the nature of their underlying representations. However, there are several ways to go about this analysis.

Notably, many prior approaches harnessed the intuition that syntactic categories allow children to transfer knowledge about how a word from one category (a NEG like *don't*) combines with words from another category (VERBs like *go* and *believe*) in order for the child to generate novel productions. For example, if the two-word combination *don't go* hasn't been heard before, then it must have been generated based on units that are more abstract than individual lexical items. This combinatory *productivity* – that is, the generation of combinations that haven't been heard before – is a sign of abstract syntactic category knowledge. Yet, what about combinations that have in fact been heard before? How do we know if children are generating them in a productive way that relies on syntactic categories (the way adults would) or instead in some other way that relies on the individual lexical items? This is where prior approaches diverge from each other. Below, we describe our approach, which is inspired primarily by Yang (2010), Yang (2011), and Pine et al. (2013).

3 Possible child category representations for multi-word combinations

We consider three types of syntactic category representation that very young children could use to form multi-word combinations (like *don't go*). The representation types differ with respect to whether the child produces multi-word combinations according to (i) the distribution of multi-word combinations in her input (NOT productive), (ii) both her input distributions and an internal category representation (SEMI-productive), or (iii) internal category representations alone (fully PRODuctive).

A child using a NOT productive representation can only generate a multi-word combination if she's heard it in her input (e.g., *don't go* $\rightarrow$ (*don't+go*)$_{Input}$). So, any multi-word combination she generates is effectively a memorized amalgam; how often she generates a particular amalgam depends on how frequently that amalgam was in her input. This contrasts with a child using a SEMI-productive representation, who relies on an internal category for generating one part of the multi-word combination and her input combinations with that category for generating the other part (e.g., *don't go* $\rightarrow$ (AUX+*go*)$_{Input}$). Here, if she's heard *go* used with an AUX – any AUX, not just *don't* – she can generate *don't go* this way. So, the child can generate some novel expressions, but still relies on input distributions when the expressions involve words that aren't part of a syntactic category. However, a child with a fully PRODuctive representation can generate novel combinations by relying on her internal syntactic categories alone, rather than input distributions of multi-word combinations (e.g., *don't go* $\rightarrow$ AUX+VERB). That is, the child draws on her internal category knowledge when generating utterances the way we believe adults typically do, and has the greatest capacity for novel multi-word combinations.

4 How can we quantitatively measure representational knowledge?

4.1 Lexical overlap as a measure of category knowledge

LEXICAL OVERLAP is often used as a measure for productivity (Yang 2010, 2011, Pine et al. 2013), and is meant to capture the intuition that words in one category can be freely combined with words from another. That is, category members are effectively interchangeable in those combinations. For example, an AUX category would allow any of its member words (e.g., *don't*, *do*, *can*, etc.) to combine with verbs like *go*. So, we would expect to see multiple auxiliaries used with any given verb (e.g., *don't go, do go, can go*, etc.) – that is, there would be overlap in the use of auxiliary lexical items. So, to assess a category, we need to examine its lexical overlap with respect to words that the category can combine with. For example, when assessing AUX, we can look at how many verbs have lexical overlap when it comes to auxiliaries.

We assess both the OBSERVED lexical overlap present in a speaker's productions and the EXPECTED lexical overlap if the speaker used a particular representation to generate those productions. While there's only one Observed score per potential category (e.g., an $\text{AUX}_{Observed}$ for how auxiliaries combine with verbs), there's an Expected score for each potential representation the speaker could be using to generate her productions. If the expected overlap for a particular representation matches the observed overlap well enough, this indicates that representation is compatible with the speaker's output.

At the category level, the two representations are that the category is (i) present (i.e. $\{$*don't, do, can*, etc.$\} \in$ AUX), or (ii) absent (i.e. *don't, do, can*, etc. are simply individual words that aren't interchangeable syntactically). At the multi-word combination level, we focus on combinations made up of two potential categories (e.g. *don't go*, which could involve AUX and VERB). For these combinations, there are three possible representations: NOT, SEMI, and fully PRODuctive. More specifically, a NOT productive representation has both categories absent; a SEMI-productive representation has one category present and the other absent; a fully PRODuctive representation has both categories present.

4.2 Calculating Observed and Expected overlap

We first describe how to calculate the lexical overlap for a potential category with respect to a set of words it combines with. This is the core calculation that will be used for calculating Observed and Expected overlap scores for multi-word combinations. We then describe how to calculate the Observed overlap for multi-word combinations and the Expected overlap for each of the three representation types (NOT, SEMI, and PROD).

For a potential category whose status is $Unknown$ (like AUX), we look at the lexical overlap in words which that category combines with (like verbs, which would be $w_{comb} \in Combine$ in (1)). Lexical overlap itself is defined very conservatively, following previous studies using it (Yang 2010, 2011, Pine et al. 2013): if more than one word $w_{unk} \in Unknown$ (e.g., both *don't* and *can*) appears in combination with a word $w_{comb} \in Combine$ (e.g., *go*), then lexical overlap for w_{comb} is 1. Otherwise, if w_{comb} only ever appears in combination with a single word $w_{unk} \in Unknown$ (e.g., *don't go* is the only combination of an auxiliary with *go*), lexical overlap is 0. This is $overlap_{w_{comb}}$ in (1). The total overlap $overlap_{Combine}$ is the lexical overlap average across all words that the potential category can combine with ($w_{comb} \in Combine$). For example, this would be the lexical overlap average across all verbs when assessing potential category AUX on how it combines with verbs. So, if there are 50 verbs that combine with auxiliaries in the data sample, then individual overlap scores $overlap_{w_{comb}}$ are calculated for each of these 50 verbs, and the average is taken of all 50 scores.

(1)
$$overlap_{w_{comb}} = \begin{cases} 1: w_{comb}\ occurs\ with\ >1\ word\ w_{unk} \in Unknown \\ 0: w_{comb}\ occurs\ with\ only\ 1\ word\ w_{unk} \in Unknown \end{cases}$$
$$overlap_{Combine} = \frac{\sum_{w_{comb} \in Combine} overlap_{w_{comb}}}{|Combine|}$$

For a multi-word combination involving two potential categories (e.g., AUX+VERB), observed overlap can be calculated with respect to each category (e.g., with respect to verbs when assessing AUX and with respect to auxiliaries when assessing VERB). The observed overlap calculation is just as in (1), shown in (2) over the set of speaker productions that involve those kind of multi-word combinations S_{Obs} (e.g., all combinations of auxiliaries+verbs for AUX+VERB).

(2)
$$Observed = overlap_{Combine}(S_{Obs})$$

Expected overlap, as mentioned, depends on the representation the speaker uses to generate her multi-word combinations. A more detailed walk-through of the Expected overlap calculation for all three representation types is in Appendix A in the supplementary materials[1], but we sketch the core intuitions here.

A child using a NOT productive representation (e.g., *don't go* $\rightarrow$ *don't+go*) generates multi-word combinations as memorized amalgams from her input, based on the frequency of those input amalgams. To simulate this, we generate multi-word combination data samples $S_{Exp_{Not}}$ that are the same size as the observed speaker multi-word combination sample S_{obs}; these samples are drawn from the speaker's input. So, if there are 100 auxiliary+verb combinations in the speaker's output, we generate 100 auxiliary+verb combinations, based

[1] Available at http://sites.uci.edu/alandibates/files/2018/07/Bates_Pearl_Braunwald_2018_BLS.pdf and www.socsci.uci.edu/~lpearl/papers/BatesPearlBraunwald2018_BLS.pdf.

on the auxiliary+verb distribution in the speaker's input.

Suppose we have a word w_{unk} (like *don't*) from a category with $Unknown$ status (like AUX). The combinations that w_{unk} is generated with depend on the combinations from the speaker's input that w_{unk} appeared with. So, the probability of sample s_i being w_{unk}+w_{comb} (e.g., *don't+go*) depends on how often w_{unk}+w_{comb} appeared in the speaker's input ($p_{w_{unk}w_{comb_{Input}}}$). We then calculate the lexical overlap of the NOT sample and use that as the Expected overlap for a child using the NOT productive representation (3).

(3)
$$s_i \in S_{Exp_{Not}}, s_i = w_{unk}w_{comb} \ \propto p_{w_{unk}w_{comb_{Input}}}$$
$$Expected_{Not} = overlap_{Combine}(S_{Exp_{Not}})$$

We can use a similar approach to calculate the Expected overlap for the SEMI-productive representation (e.g., *don't go* → AUX+*go* or *don't*+VERB). For simplicity, we abbreviate the word from the category as w_{+cat} and the word not from a category as w_{-cat}. Then, to generate combination $w_{unk}w_{comb}$, the child relies on her internal category representation to generate word w_{+cat} and looks to her input to see how often words from this category combine with word w_{-cat}. So, she would generate combination $w_{unk}w_{comb}$ with about the same frequency she heard examples of either $Unknown$+w_{comb} (if $Unknown$ is the category) or w_{unk}+$Combine$ (if $Combine$ is the category). To simulate this process, we generate multi-word combination data samples $S_{Exp_{Semi}}$ that are the same size as the observed speaker multi-word combination sample S_{obs}. The probability of multi-word sample $s_i \in S_{Exp_{Semi}}$ involving a specific word $w_{+cat} \in Category$ combined with w_{-cat} depends on how often any word in $Category$ combines with w_{-cat} in the speaker's input ($p_{Category}\ p_{w_{-cat_{Input}}}$). We then calculate the lexical overlap for the SEMI sample and use that as the Expected overlap for a child using a SEMI-productive representation (4).

(4)
$$s_i \in S_{Exp_{Semi}}, s_i = w_{+cat}w_{-cat} \ \propto p_{Category}\ p_{w_{-cat_{Input}}}$$
$$Expected_{Semi} = overlap_{Combine}(S_{Exp_{Semi}})$$

A child with a fully PRODuctive representation (e.g., *don't go* → AUX+VERB) generates her multi-word combinations by relying on internal category representations for both words. Yang (2010, 2011) describes an analytical solution for the Expected lexical overlap when both categories exist (5). We can use this here, rather than generating expected samples and calculating lexical overlap for those samples. The key intuition involves the definition of lexical overlap, where a word w_{comb} shows lexical overlap if more than one word $w_{unk} \in Unknown$ combines with w_{comb}. So, we can calculate this analytically as 1 minus the probability that w_{comb} will (i) never appear with any word in $Unknown$, or (ii) only appear with a single word in $Unknown$. This is equivalent to the formula in (5) for the Expected overlap for word w_{comb}, whose derivation is discussed more fully in Appendix A of the Supplementary Materials. All word probabilities are estimated based on the speaker's productions of w_{comb} and w_{unk} (i.e., $p_{w_{comb}} = p_{w_{comb_{Obs}}}$, $p_{w_{unk}} = p_{w_{unk_{Obs}}}$). This is because all words in these combinations are generated from an underlying internal category, and so don't rely on the speaker's input. As with the original calculation of lexical overlap, these individual word overlaps are averaged to get the Expected overlap.

(5)
$$\begin{aligned} overlapprod_{w_{comb}} &= 1 - P(no\; w_{comb}) - P(only\; 1\; w_{comb}) \\ &= 1 + (|Unknown| - 1)(1 - p_{w_{comb}})^{S_{obs}} \\ &- \sum_{w_{unk} \in Unknown} (p_{w_{comb}} * p_{w_{unk}} + 1 - p_{w_{comb}})^{S_{obs}} \end{aligned}$$
$$Expected_{Prod} = \frac{\sum_{w_{comb} \in Combine} overlapprod_{w_{comb}}}{|Combine|}$$

5 Data

Given that syntactic category knowledge may be present as early as two years old, we investigated data from a child (hereafter **L**) just before the age of two. L's productions between 20 and 24 months were hand-recorded in daily diary data in the Susan R. Braunwald Language Acquisition Diaries (Braunwald 2015), and represent a rich cross-contextual sample of L's whole language acquisition experience. We also included child-directed mealtime input from L's caretaker when L is between 20 and 24 months, which are in the Braunwald corpus (Braunwald 1995) in CHILDES (MacWhinney 2000). Between these two datasets, we had a dense longitudinal sample from the same child of both her output (from the daily diary data) and her input (from the child-directed mealtime speech).

Because verbs are often considered the backbone of language, given the wealth of information they encode about events and their participants (Gleitman 1990, Tomasello & Merriman 1995), we focused our investigation on syntactic categories in verb phrases (VPs): VERB itself, along with NEG, AUX, ADJective, PREPosition, and NOUN. The VPs from L and her caretaker were manually extracted and syntactically annotated, yielding 2,154 child-produced VPs from L and 2,184 adult-produced VPs from L's caretaker.[2] We additionally restricted our analyses to lexical items shared by L and her caretakers to facilitate comparisons between their lexical overlaps, as Pine et al. (2013) note that differing vocabulary sizes can disrupt comparisons between subjects. This yielded 105 verbs total. Because the quantitative analysis we use requires sample sizes that are sufficiently large, we decided to only include potential categories where there were at least 100 tokens in L's productions (Goldin-Meadow & Yang (2016), Charles Yang, p.c.). For example, at least 100 instances of nouns combining with verbs were needed to include the potential category NOUN. This led to us including two open-class categories (VERB and NOUN) and two closed-class categories (AUX and NEG). Table 1 shows the types and tokens from L and L's caretaker for each potential category and multi-word combination involving those potential categories.[3]

6 Evaluating the possible representations

Because we consider four potential categories in VPs, a child's complete category representation involves something about each potential category. In particular, for each of the four categories, the child either has a category representation for it (e.g., all verbs categorized as VERB) or doesn't (e.g., all verbs treated as individual words). This yields 16 possible category representations (2^4) the child might have. The completely NOT productive representation (Rep_{NOT}) has no categories – that is, all four potential categories are absent, and the words that would be in them are represented only as individual words. In contrast, the fully PRODuctive representation (Rep_{PROD}) is the one where all four poten-

[2]We note that L's verb usage seems typical of her age group, as assessed by a corpus analysis of verb production frequency from 93 North American English children between the ages of 20 and 24 months from the CHILDES database (MacWhinney 2000). In particular, based on L's verbs and the verbs used by these 93 children (10432 verb tokens and 322 verb types), L used 16 of their 20 most frequent verbs, and they collectively used 15 of her 20 most frequent verbs.

[3]We note that the multi-word combination counts were analyzed irrespective of order. So, for example, VERB+NOUN combinations include instances such as *I go* (NOUN VERB) and *have coffee* (VERB NOUN). We also note that contractions (e.g., *don't*) were analyzed as belonging to both potential categories their components were from. For example, *don't* was counted as an instance of both an AUX and a NEG, because *do* is an AUX for adults and *n't* is a NEG for adults.

Table 1: Types and tokens of potential categories and multi-word combinations involving those categories in the VPs of L and L's caretaker.

	L		**L's caretaker**	
Potential category	Types	Tokens	Types	Tokens
VERB	105	2642	105	3164
NOUN	504	2330	617	2606
AUX	21	198	38	454
NEG	6	114	11	104
Multi-word combination	Types	Tokens	Types	Tokens
VERB+NOUN	1111	2330	1426	2606
VERB+AUX	95	198	239	454
VERB+NEG	42	114	61	104

tial categories are present for the child, just as they are for adults like L's caregiver. The remaining possible category representations are SEMI-productive (Rep_{SEMI}), because they involve at least one category present and at least one absent. For example, one possible Rep_{SEMI} would have VERB and NEG as categories while nouns and auxiliaries would be represented as individual words only.

We can evaluate each category representation based on how well its Expected lexical overlap matches L's Observed lexical overlap. Because we have four potential categories (VERB, NOUN, AUX, NEG), we can calculate lexical overlap scores for all multi-word combinations involving these potential categories. More specifically, for any multi-word combination, we calculate lexical overlap scores with respect to the *Unknown* category, and either word in the multi-word combination can be the *Unknown* category being assessed. For example, in VERB+NOUN multi-word combinations, either the VERB or the NOUN could be assessed as the *Unknown* category while the other category serves as the collection of words in *Combine* (i.e., *Unknown*=NOUN and *Combine*=VERB, or vice versa). The Observed and Expected lexical overlap scores are calculated based on which set of words is *Unknown* and which set is *Combine*.

Importantly, the Expected calculation depends on the status of *Unknown* and *Combine* in the category representation itself. For example, consider *Unknown*=NOUN while *Combine*=VERB. A representation where both categories were absent would calculate the Expected overlap score using $Expected_{Not}$, a representation where NOUN was present while VERB was absent would use $Expected_{Semi}$, and a representation where both categories were present would use $Expected_{Prod}$. A more detailed walk-through of this calculation is in Appendix A of the Supplementary Materials.

Once we have the Observed and Expected lexical overlap scores for a category representation, how do we tell that they match sufficiently? Following Goldin-Meadow & Yang (2016), we use Lin's Concordance Correlation Coefficient (**LCCC**, represented with ρ_c: Lawrence & Lin (1989)) to assess agreement between the Observed and Expected overlap. LCCC measures the agreement between two sets of observations on a scale from -1 to 1, with a ρ_c of -1 indicating perfect disagreement, 1 indicating perfect agreement, and 0 indicating no agreement. So, given that there are multiple lexical overlap scores for each category representation (one for each legitimate multi-word combination within a particular category representation), we assess ρ_c for the Observed vs. Expected overlap scores within that category representation (Table 2).

With ρ_c scores for each of the 16 possible category representations, we then need to decide which representations have a "good enough" match between Observed and Expected overlap. Unfortunately, there isn't a current consensus about what the threshold should be for good agreement with the LCCC (Altman 1990, McBride 2005). Given this, we decided to leverage L's input data, with the idea that L's caregiver had a fully productive category representation (Rep_{PROD}) involving VERB, NOUN, AUX, and NEG. Because of this, the agreement between the Observed overlap in L's caregiver's productions and the Expected overlap from the Rep_{PROD} category representation could serve as a "good enough" threshold of agreement. More specifically, because we believe the Rep_{PROD} category representation generated L's caregiver's productions, the ρ_c obtained for that representation is a reasonable cutoff for when a category representation in general matches sufficiently well with the observed data. We found ρ_c = 0.901 when comparing the Expected overlap from a Rep_{PROD} category representation against the Observed overlap in L's caretaker's productions. We take this value as our threshold for when L's possible category representations are sufficiently compatible with her output (Table 2).

Though agreement values range from ρ_c = 0.715 to 0.935, only two of the sixteen possible category representations are above the "good enough" threshold of 0.901. Both

Table 2: LCCC scores for the 16 possible category representations L could have, comparing her Observed lexical overlap against the lexical overlap Expected by each possible category representation. Representations with sufficient agreement (>0.901) are indicated.

Represention	VERB	NOUN	AUX	NEG	***LCCC*** ρ_c	**Sufficient agreement?**
Rep_{NOT}	✗	✗	✗	✗	0.873	
$\text{Rep}_{\text{PRODUCTIVE}}$	✓	✓	✓	✓	0.838	
$\text{Rep}_{\text{SEMI}_1}$	✓	✓	✓	✗	0.851	
$\text{Rep}_{\text{SEMI}_2}$	✓	✓	✗	✓	0.802	
$\text{Rep}_{\text{SEMI}_3}$	✓	✗	✓	✓	0.753	
$\text{Rep}_{\text{SEMI}_4}$	✗	✓	✓	✓	0.867	
$\text{Rep}_{\text{SEMI}_5}$	✓	✓	✗	✗	0.809	
$\text{Rep}_{\text{SEMI}_6}$	✓	✗	✓	✗	0.753	
$\text{Rep}_{\text{SEMI}_7}$	✓	✗	✗	✓	0.719	
$\text{Rep}_{\text{SEMI}_8}$	✗	✗	✓	✓	0.915	Yes
$\text{Rep}_{\text{SEMI}_9}$	✗	✓	✗	✓	0.891	
$\text{Rep}_{\text{SEMI}_{10}}$	✗	✓	✓	✗	0.765	
$\text{Rep}_{\text{SEMI}_{11}}$	✓	✗	✗	✗	0.715	
$\text{Rep}_{\text{SEMI}_{12}}$	✗	✓	✗	✗	0.794	
$\text{Rep}_{\text{SEMI}_{13}}$	✗	✗	✓	✗	0.850	
$\text{Rep}_{\text{SEMI}_{14}}$	✗	✗	✗	✓	0.935	Yes

category representations are SEMI-productive ($\text{Rep}_{\text{SEMI}_8}$=0.915, $\text{Rep}_{\text{SEMI}_{14}}$=0.935). As Table 2 shows, neither category representation involves knowledge of the open-class categories VERB or NOUN. Instead, both involve knowledge of the closed-class category NEG (covering all 6 negations in L's productions) and one also includes knowledge of AUX (covering all 21 auxiliaries in L's productions).

7 General discussion

We find quantitative evidence for adult-like closed-category knowledge in a child before age two, but not for adult-like open-class category knowledge. This suggests open-class categories may take longer to develop into adult-like categories, compared with closed-class categories. Notably, this supports advocates of early closed-class category knowledge (Capdevila i Batet & Llinàs i Grau 1995, Rowland & Theakston 2009, Theakston

& Rowland 2009, Yang 2010, Shin 2012, Meylan et al. 2017). This also aligns with the idea that closed-class categories may provide a way of breaking into the categorization of open-class lexical items (Höhle et al. 2004).

One reason why adult-like closed-class categories might emerge sooner is that they have fewer members than their open-class counterparts. For instance, in our data sample, there were 6 negations and 21 auxiliaries, in contrast with 105 verbs and 504 nouns. So, with fewer members, it may be easier to cluster all relevant lexical items into their respective closed-class categories.

Another difference is the existence of salient semantic sub-classes within the open-class categories – these sub-classes might form a natural category for the child, rather than the child clustering all relevant lexical items into an adult-like NOUN or VERB. For example, 6-month-olds recognize concrete nouns specifically (Bergelson & Swingley 2012), and so this kind of noun might persist as a natural class even after children recognize and use other nouns. Similarly, 3- and 4-year-olds have distinct comprehension behavior for passives involving actional verbs like *hug* vs. non-actional verbs like *surprise, find, forget*, or *love* (Nguyen et al. 2016, Nguyen & Pearl 2018). So, the actional lexical feature (and others) may cause younger children to form categories for subsets of verbs, rather than having a single VERB category.

The existence of these potential child categories that are subsets of the adult category highlights an interesting area for future research: evaluate potential child-like categories against young children's productions, such as L's data. Recall that we only evaluated adult-like categories here, which encompass all relevant lexical items (e.g. all nouns for NOUN). However, the very same quantitative approach can be used to assess whether potential child-like categories, which may be subcategories of the adult versions (e.g., CONCRETE-NOUN vs. NON-CONCRETE-NOUN), best match children's productions. All that's needed is to define which words belong to which child-like category, and the possible multi-word combination types involving these child-like categories. If a good enough match were found to the child's productions, this would provide quantitative support for a specific child-like category, distinct from the absence of any category or from the presence of an adult-like category. That is, we would have quantitative support for a particular developing category. We leave this exciting possibility for the future.

REFERENCES

ALTMAN, DOUGLAS G. 1990. *Practical statistics for medical research.* CRC press.

BERGELSON, ELIKA, and DANIEL SWINGLEY. 2012. At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences* 109. 3253–3258.

BOOTH, AMY, and SANDRA WAXMAN. 2003. Mapping words to the world in infancy: On the evolution of expectations for nouns and adjectives. *Journal of Cognition and Development* 4. 357–381.

BRAUNWALD, SUSAN. 1995. *Differences in the acquisition of early verbs: Evidence from diary data from sisters*. 81–111. Psychology Press.

BRAUNWALD, SUSAN. 2015. Susan R. Braunwald Language Acquisition Diaries. In *UC Irvine Special Collections and Archives*. UC Irvine.

CAPDEVILA I BATET, MONTSERRAT, and MIREIA LLINÀS I GRAU. 1995. The acquisition of negation in English. *Atlantis* 27–44.

GLEITMAN, LILA. 1990. The structural sources of verb meanings. *Language Acquisition* 1. 3–55.

GOLDIN-MEADOW, SUSAN, and CHARLES YANG. 2016. Statistical evidence that a child can create a combinatorial linguistic system without external linguistic input: Implications for language evolution. *Neuroscience & Biobehavioral Reviews* .

HÖHLE, BARBARA; JÜRGEN WEISSENBORN; DOROTHEA KIEFER; ANTJE SCHULZ; and MICHAELA SCHMITZ. 2004. Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy* 5. 341–353.

KEMP, NENAGH; ELENA LIEVEN; and MICHAEL TOMASELLO. 2005. Young children's knowledge of the determiner and adjective categories. *Journal of Speech, Language, and Hearing Research* 48. 592–609.

LAWRENCE, I., and KUEI LIN. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 255–268.

MACWHINNEY, B. 2000. The CHILDES project: The database. Psychology Press.

MCBRIDE, G.B.. 2005. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. *NIWA Client Report: HAM2005-062* .

MEYLAN, STEPHAN C; MICHAEL C FRANK; BRANDON C ROY; and ROGER LEVY. 2017. The emergence of an abstract grammatical category in children's early speech. *Psychological Science* 28. 181–192.

NGUYEN, EMMA; DIANE LILLO-MARTIN; and WILLIAM SNYDER. 2016. Actionality speaks louder than felicity: Children's comprehension of long passives. In *Proceedings from the Generative Approaches to Language Acquisition*. 232–246.

NGUYEN, EMMA, and LISA PEARL. 2018. Do You Really Mean It? Linking Lexical Semantic Profiles and the Age of Acquisition for the English Passive. In *35th West Coast Conference on Formal Linguistics*. 288–295. Cascadilla Proceedings Project.

PINE, JULIAN; DANIEL FREUDENTHAL; GRZEGORZ KRAJEWSKI; and FERNAND GOBET. 2013. Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition* 127. 345–360.

PINE, JULIAN, and ELENA LIEVEN. 1997. Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics* 18. 123–138.

PINKER, STEVEN. 1984. *Language learnability and language development.* Cambridge, MA: Harvard University Press.

ROWLAND, CAROLINE F., and ANNA L. THEAKSTON. 2009. The acquisition of auxiliary syntax: A longitudinal elicitation study. Part 2: The modals and auxiliary DO. *Journal of Speech, Language, and Hearing Research* 52. 1471–1492.

SHIN, YU KYOUNG. 2012. A New Look at Determiners in Early Grammar: Phrasal Quantifiers. *Language Research* 48. 573–608.

THEAKSTON, ANNA L.; PAUL IBBOTSON; DANIEL FREUDENTHAL; ELENA LIEVEN; and MICHAEL TOMASELLO. 2015. Productivity of noun slots in verb frames. *Cognitive Science* 39. 1369–1395.

THEAKSTON, ANNA L., and CAROLINE F. ROWLAND. 2009. The acquisition of auxiliary syntax: A longitudinal elicitation study. part 1: Auxiliary BE. *Journal of Speech, Language, and Hearing Research* 52. 1449–1470.

TOMASELLO, MICHAEL. 2004. What kind of evidence could refute the UG hypothesis? *Studies in Language* 28. 642–645.

TOMASELLO, MICHAEL, and SILKE BRANDT. 2009. Flexibility in the semantics and syntax of children's early verb use. *Monographs of the Society for Research in Child Development* 74. 113–126.

TOMASELLO, MICHAEL, and WILLIAM E. MERRIMAN. 1995. *Beyond names for things: Young children's acquisition of verbs*. Psychology Press.

VALIAN, VIRGINIA. 1986. Syntactic categories in the speech of young children. *Developmental Psychology* 22. 562.

YANG, CHARLES. 2010. Who's Afraid of George Kingsley Zipf. Unpublished Manuscript. Online: https://www.ling.upenn.edu/ ycharles/papers/zipfnew.pdf.

YANG, CHARLES. 2011. A statistical test for grammar. In *Proceedings of the 2nd workshop on Cognitive Modeling and Computational Linguistics*. 30–38. Association for Computational Linguistics.

Quantitative Comparison for Generative Theories: Embedding Competence Linguistic Theories in Cognitive Architectures and Bayesian Models

ADRIAN BRASOVEANU[†] & JAKUB DOTLAČIL[‡]
[†]*UC Santa Cruz* & [‡]*University of Amsterdam* [*]

1 Introduction

The main goal of this paper is to introduce a framework integrating generative theories, computational cognitive models, and Bayesian methods. The integration proceeds in two parts. First, competence-level generative theories are embedded in performance-level processing theories formulated in the ACT-R cognitive architecture (Adaptive Control of Thought-Rational; Anderson and Lebiere 1998, Lewis and Vasishth 2005 a.o.). Second, these integrated competence-performance processing theories become part of a Bayesian model, which can be fitted to experimental data.

The main upshot is that we are able to consider alternative generative grammar theories and quantitatively compare how well they fit experimental data. A detailed introduction to the framework will be available soon in Brasoveanu and Dotlačil 2019. In this paper, we focus on a case study: the lexical decision task in Murray and Forster 2004. We model their data with 3 different ACT-R models that differ qualitatively and/or quantitatively. We then fit these models to the Murray and Forster 2004 experimental data and compare the results.

Our generative grammar + ACT-R + Bayes framework is very general: it enables us to incorporate rich syntactic and semantic theories, and also model experimental tasks other than the one considered here. We choose to model lexical decision for reasons of clarity and transparency: this task is straightforward, so it makes the general structure of our approach obvious. The integration of generative theories and ACT-R is computationally implemented in a new Python3 library *pyactr*.[1]

[*]We are grateful to Donka Farkas, Abel Rodriguez, Matt Wagers and the UCSC S-lab audience (January 2018) for comments and discussion. The usual disclaimers apply.

[1]Readers familiar with ACT-R know that the official implementation of ACT-R is in Lisp. Using *pyactr* enables us to easily interface ACT-R models with widely-used Python3 libraries for Bayesian modeling, for example, *pymc3*. The most relevant parts of the code are provided in the main text; the full code will be

The paper is structured as follows. In Section §2, we introduce the lexical decision task and the data we model, and discuss a basic Bayesian log-frequency model for this data. This model highlights the imperfect data fit of the log-frequency assumption, and introduces the basic structure of Bayesian models that we will need later. In Section §3, we introduce the main idea behind our ACT-R models of lexical decision: frequency effects as practiced memory retrieval. In Section §4, we introduce a series of 3 ACT-R models of a participant completing the lexical decision task, and we quantitatively compare them. These lexical access models are particularly simple; the concluding section (§5) briefly outlines how the framework can accommodate much more realistic linguistic theories.

2 The lexical decision task and a Bayesian log-frequency model

Word frequency is one very robust parameter affecting latencies and accuracies in lexical decision tasks (Whaley, 1978). Frequency effects have been found in many, if not all tasks that involve some kind of lexical processing (Forster, 1990; Monsell, 1991). These effects are assumed to have a specific functional form: lexical access latency can be well approximated as a log-function of word frequency (Howes and Solomon, 1951).

Murray and Forster 2004 studied the role of frequency in detail and identified various issues with the log-frequency model. Their data consisted of collected responses and response times in a lexical decision task using words from 16 frequency bands: these 16 word-frequency bands (measured in tokens-per-million) together with the results of Experiment 1 in Murray and Forster 2004, are provided in Table 1 below. In what follows, we will use the mean frequency listed in the second column from the left as the predictor variable, and we will model the lexical decision latencies and accuracies reported in columns three and four in terms of this predictor.

To get acquainted with the structure of a Bayesian model, and as a baseline for all our future models, we specify a simple Bayesian log-frequency model for this data. The basic structure of the model is provided in Table 2. The last line in Table 2 shows that the log-frequency model takes the observed reactions times (RTs) for a word to be a linear function of the log-frequency of that word + some normal/Gaussian-distributed noise. This

available in Brasoveanu and Dotlačil 2019.

Table 1: Exp. 1 in Murray and Forster (2004)

Frequency range	**Mean frequency**	**Latency (ms)**	**Accuracy (%)**
315–197	**242.0**	**542**	**97.22**
100–85	**92.8**	**555**	**95.56**
60–55	**57.7**	**566**	**95.56**
42–39	**40.5**	**562**	**96.3**
32–30	**30.6**	**570**	**96.11**
24–23	**23.4**	**569**	**94.26**
19	**19.0**	**577**	**95**
16	**16.0**	**587**	**92.41**
14-13	**13.4**	**592**	**91.67**
12–11	**11.5**	**605**	**93.52**
10	**10.0**	**603**	**91.85**
9	**9.0**	**575**	**93.52**
7	**7.0**	**620**	**91.48**
5	**5.0**	**607**	**90.93**
3	**3.0**	**622**	**84.44**
1	**1.0**	**674**	**74.63**

linear function is the LIKELIHOOD component of our Bayesian model, which connects the model to the data.

To estimate a linear function, we need to estimate two parameters: its intercept and its slope. A Bayesian model specifies our prior beliefs about these parameters (the first three rows in Table 2), which can be very vague and unconstrained, for example, we can take them to be normally distributed with a mean of 0 and a large standard deviation, for example, 300 ms.[2] Furthermore, we also specify a half-normal prior distribution for the noise – half-normal because the noise parameter is a standard deviation, so it is has to be positive. The Bayesian model then updates these priors with the information provided by the data, and outputs the posterior distributions of these parameters; technically speaking, the model draws sufficiently many samples from the posterior distributions such that the resulting empirical distributions approximate the true posteriors very well (the last row in Table 2).

The full code for all the models introduced in this paper, as well as detailed discus-

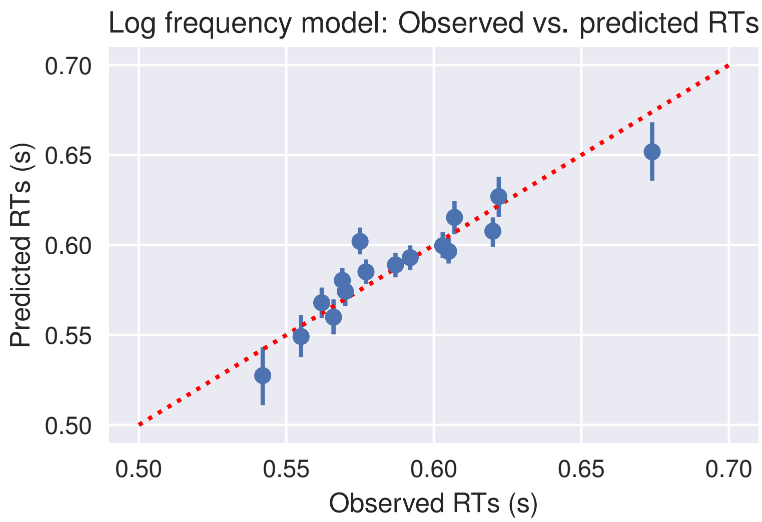
[2]The (half-)Gaussians in Table 2 are parametrized in terms of their means and variances, as is customary.

Table 2: Bayesian log-frequency model

Priors		
intercept	$\sim$	Normal(0, 300^2)
slope	$\sim$	Normal(0, 300^2)
noise	$\sim$	HalfNormal(0, 300^2)
Likelihood		
RT	$\sim$	Normal(intercept + slope $\cdot$ log(freq), noise^2)

sions of related technical issues, are provided in Brasoveanu and Dotlačil 2019. For our purposes, it is enough to examine the plot in Figure 1 of the posterior predictions made by the log-frequency model. On the x axis, we plot the observed RTs (in seconds) for the 16 word frequency bands – see the blue dots, which correspond to column 3 in Table 1 above. On the y axis, we plot the mean posterior RTs in seconds (the same blue dots) and the associated 95% credible intervals (CRIs). The diagonal red line helps visualize the fit of the model to the data: the closer the blue dots are to the line, the closer the estimated RTs are to the observed data, and the better the model is at capturing the data. We see that the log-frequency model gets middle values right, but underestimates the time needed to access words in extreme frequency bands.

Figure 1: Posterior predictions of log-frequency model



3 Frequency effects as practiced memory retrieval in ACT-R

Our proposal is to model frequency effects as practiced memory retrieval: latency in memory recall is a power function of the amount of practice, and also of the time elapsed since individual instances of practice (Anderson 1982, Logan 1990, Anderson, Fincham, et al. 1999). By practice, we simply mean the repeated presentation of an item, for example, the repeated exposure to and/or use of a word in daily conversation.[3]

A concrete implementation of practiced memory retrieval is provided in the ACT-R cognitive architecture. The (base) activation of an item i is A_i. It is based on the amount of time t_k elapsed since each rehearsal k of a word, and it is a power function of time because of its form t_k^{-d}. Activations contributed by individual rehearsals k (from 1 to the total number n of word rehearsals) are summed, and the total sum is log-compressed – see the formula in (1) below. The activation of an item is in turn used to compute accuracy (2) and latency (3) for retrieval processes. The free parameters associated with each formula are boxed in the equations and enumerated in parentheses.

(1) $$A_i = \log\left(\sum_{k=1}^{n} t_k^{-\boxed{d}}\right)$$ ($\boxed{d}$: decay)

(2) $$P_i = \left[1 + \exp\left(-\frac{A_i - \boxed{\tau}}{\boxed{s}}\right)\right]^{-1}$$ ($\boxed{s}$: noise, $\boxed{\tau}$: threshold)

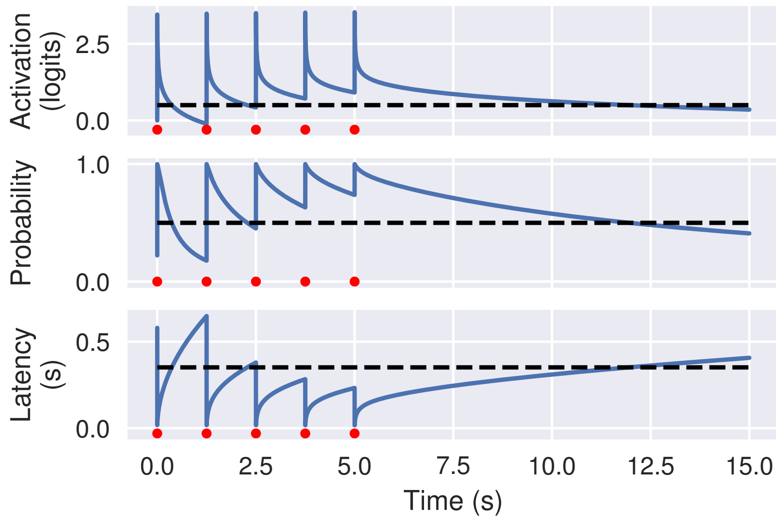
(3) $$T_i = \boxed{F} e^{-\boxed{f} A_i}$$ ($\boxed{F}$: latency factor, $\boxed{f}$: latency exponent)

As an example, Figure 2 plots activation, retrieval probability and retrieval latency as a function of time for an item presented five times at equally spaced 1.25 s intervals. The top plot shows that the activation of an item sharply increases after every presentation, and then decays until the next presentation. Importantly, after every presentation, the decay curve becomes shallower, ensuring that the item stays activated, that is, above the threshold (the dotted black line), for a longer period of time. The second plot shows that the probability of successfully retrieving an item closely follows its activation. Finally, the third plot shows that latency of retrieval is inversely related to the activation of an item:

[3] This proposal is different from the one in Murray and Forster 2004.

the higher the activation of an item, the less time it takes to retrieve it.

Figure 2: Activation, retrieval probability, and retrieval latency as a function of time (threshold – dotted black line; 5 presentations – red)



How do we estimate the schedule of presentations for words? For any word, the number of rehearsals that contribute to its activation are determined by its frequency (we ignore other factors throughout this paper). We generate a presentation schedule for a 15-year old speaker based on word frequency and the average number of words the 15-year old speaker is estimated to have seen (estimate based on Hart and Risley 1995; see Brasoveanu and Dotlačil 2019 for more details). With this schedule in place, we can compute activations for all 16 word-frequency bands and store them in a 16-coordinate vector we will call ACTIVATION-FROM-TIME.

A Bayesian model for the lexical decision data is specified in Table 1, with ACT-R likelihoods for both lexical decision RTs and lexical decision accuracies. Embedding ACT-R models in Bayesian models enables us to link them to experimental data (in the case at hand: the lexical decision data from Murray and Forster 2004), and to estimate the parameters of the ACT-R models based on that data.

Table 3 shows how we specify this type of ACT-R + Bayes models. Just as in the simpler log-frequency model, we have vague, low information priors for the parameters

Table 3: ACT-R + Bayes models of lexical decision

Priors		
d	$\sim$	Uniform(0, 1)
s	$\sim$	Uniform(0, 5)
τ	$\sim$	Normal(0, 10^2)
F	$\sim$	HalfNormal(0, 1^2)
f	$\sim$	HalfNormal(0, 1^2)
Likelihood		
μ_{RT}	$=$	PYACTR-MODEL(ACTIVATION-FROM-TIME, d, F, f)
μ_{PROB}	$=$	$\left[1 + \exp\left(-\frac{\text{ACTIVATION-FROM-TIME} - \tau}{s}\right)\right]^{-1}$
RT	$\sim$	Normal(μ_{RT}, 0.01^2)
accuracy	$\sim$	Normal(μ_{PROB}, 0.01^2)

of interest (the first five rows in Table 3). The function PYACTR-MODEL in the likelihood part invokes an ACT-R model implemented in *pyactr* and runs it to generate mean latencies μ_{RT} for the 16 word-frequency bands in Murray and Forster 2004. This function is parametrized by the activations for the 16 word-frequency bands ACTIVATION-FROM-TIME, the decay parameter d, the latency factor F and the latency exponent f. For simplicity, we compute mean accuracies μ_{PROB} for the 16 word-frequency bands directly using equation (2), but we can also obtain them by repeated runs of the same *pyactr* model we use to obtain latencies.

In the last two lines of Table 3, the 16 observed mean RTs and accuracies from Murray and Forster 2004 are assumed to be noisy realizations of the ACT-R generated RTs and accuracies. To see if ACT-R can precisely fit the observed values, we require the normally-distributed noise to be very small in both cases.[4]

With this framework in place, we can now fit a variety of ACT-R models to data and compare their fit by specifying the models in *pyactr* and then plugging them into the likelihood function 'slot' of our Bayesian model via the PYACTR-MODEL function.

[4]See the 0.01^2 variances in the last two lines of Table 3. These small variances, as well the normal distributions, are independently motivated by the fact that we model observed MEAN RTs.

4 Three ACT-R models of lexical decision

In this section, we discuss three ACT-R models that we fit to data by embedding them in Bayesian models via the procedure shown in Table 3. ACT-R models provide the essential link to competence-level generative theories: they embed competence theories in processing models.

More specifically, for the lexical decision task we are modeling here, we have: (*i*) a symbolic competence theory of the lexicon – the structure of a lexical entry, what information is stored in it, and so on; we oversimplify here and assume lexical entries only store the written form and syntactic category of a word; and (*ii*) a symbolic performance theory of what human participants actually do in a lexical decision task – lexical items are stored in declarative memory and have an activation that is a function of their frequency, participants read a written form (sequence of characters) on the screen and attempt to retrieve a word with that form, and so on.

These symbolic components are implemented in ACT-R as **condition-action** pairs (production rules) stored in procedural memory. These rules trigger a cognitive **action** if the cognitive context, that is, the mental state of the ACT-R mind, satisfies a range of **conditions**. Depending on which production rules we use and how we formulate them, we implement different symbolic competence and performance theories in ACT-R, which can then be quantitatively compared by fitting them to the same experimental data.

The three models we consider for the remainder of this section differ in various ways, both qualitatively (symbolically) and quantitatively (subsymbolically). For presentational simplicity, we consider symbolic and subsymbolic differences in the performance / processing hypotheses we entertain, but different competence-level representational assumptions can be implemented and compared in the exact same way.

4.1 Model 1

The first model of lexical decision we consider consists of 4 central rules. The first rule is the *attend word* rule below that takes a visual location encoded in the visual location buffer, a.k.a., the visual *where* buffer, and issues a command to the visual *what* buffer to move attention to that visual location.[5]

(4) Rule 1 (ATTEND WORD):

conditions				actions		
=goal>	STATE:	attend		=goal>	STATE:	retrieving
=visual-location>	ISA:	-location	⟹	+visual>	CMD:	move-attention
					SCREEN-POS:	=visual-location
?visual>	STATE:	free				

Rule 2 takes the visual value discovered at that visual location, which is a potential word form, and places a declarative memory request to retrieve a word with that form.

(5) Rule 2 (RETRIEVING):

conditions				actions		
=goal>	STATE:	retrieving		=goal>	STATE:	retrieval-done
			⟹			
=visual>	VALUE:	=val		+retrieval>	ISA:	word
					FORM:	=val

Rules 3 (LEXEME RETRIEVED) and 4 (NO LEXEME FOUND) take care of the two possible outcomes of the memory retrieval request: if a word with that form is retrieved from memory (LEXEME RETRIEVED), a command is issued to the motor module to press the J key, which is the Yes response; if no word is retrieved (NO LEXEME FOUND), a command is issued to the motor module to press the F key, which is the No response.

(6) Rule 3 (LEXEME RETRIEVED):

conditions				actions		
=goal>	STATE:	retrieval-done		=goal>	STATE:	done
			⟹			
?retrieval>	BUFFER:	full		+manual>	CMD:	press-key
	STATE:	free			KEY:	J

[5] For more details about the modular structure of an ACT-R mind and the structure of the peripheral modules (visual and motor) we assume here, see Brasoveanu and Dotlačil 2019.

(7) Rule 4 (NO LEXEME FOUND):

	conditions				actions	
=goal>	STATE:	retrieval-done		=goal>	STATE:	done
			$\Longrightarrow$			
?retrieval>	BUFFER:	empty		+manual>	CMD:	press-key
	STATE:	error			KEY:	F

Running this model with the string *elephant* as a stimulus displayed in the center of the screen, we obtain the temporal trace of the lexical-decision cognitive process in Table 4.

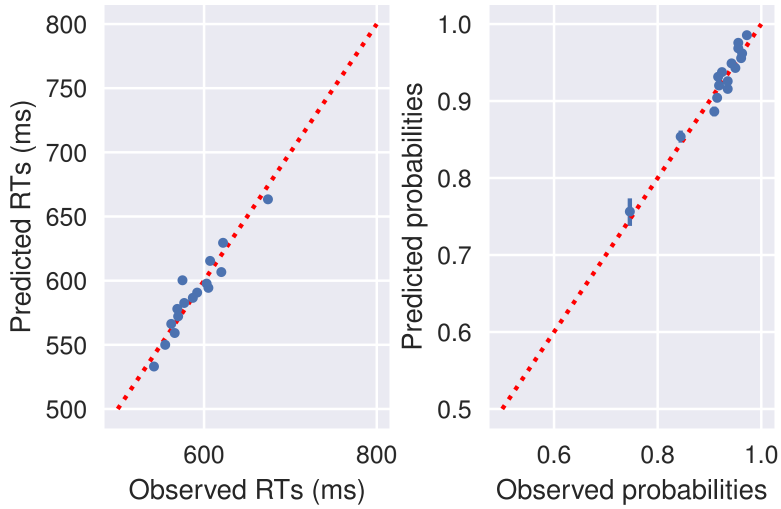
Table 4: Model 1: Temporal trace

```
****Environment: {1: {'text': 'elephant', 'position': (320, 180)}}
(0, 'PROCEDURAL', 'RULE SELECTED: attend word')
(0.05, 'PROCEDURAL', 'RULE FIRED: attend word')
(0.0679, 'PROCEDURAL', 'RULE SELECTED: retrieving')
(0.1179, 'PROCEDURAL', 'RULE FIRED: retrieving')
(0.1179, 'retrieval', 'START RETRIEVAL')
(0.1679, 'retrieval', 'RETRIEVED: word(form= elephant)')
(0.1679, 'PROCEDURAL', 'RULE SELECTED: lexeme retrieved')
(0.2179, 'PROCEDURAL', 'RULE FIRED: lexeme retrieved')
(0.2179, 'manual', 'COMMAND: press_key')
(0.4679, 'manual', 'PREPARATION COMPLETE')
(0.5179, 'manual', 'INITIATION COMPLETE')
(0.6179, 'manual', 'KEY PRESSED: J')
```

We can then take the time between the point at which a stimulus is displayed on the screen (in the 'environment') and the time at which a key is pressed as the RT that we need to fit to the experimental data from Murray and Forster 2004.[6] The posterior predictions obtained by embedding Model 1 in a Bayesian model and fitting it to data are provided in Figure 3. We see that the model fits both the latency and the accuracy data very well.

[6]We could also match the accuracy of the model (how often it presses the J key for existing words) to the Murray and Forster 2004 data by repeatedly running the *pyactr* model. As we mentioned above, for simplicity, we model accuracies directly in the Bayesian model using the ACT-R equation in (2).

Figure 3: Model 1: estimated and observed RTs and probabilities



4.2 Model 2: adding the imaginal buffer

Model 1 oversimplifies the process of encoding visually retrieved data: it assumes the visual value found at a particular visual location is immediately shuttled to the retrieval buffer. But cognition in ACT-R is GOAL-DRIVEN: any important step in a cognitive process should involve the GOAL buffer or the IMAGINAL buffer, which is a goal-like buffer storing internal snapshots of the current cognitive state. In our case, it is natural to assume that the transfer between the visual and the retrieval buffer is mediated by the IMAGINAL buffer.

We correct this oversimplification in a second model. The Bayesian model remains the same, the only part we change is the *pyactr*-provided likelihood for latencies. In particular, we modify the procedural core of the ACT-R model. First, we add the imaginal buffer to the model. Then, we replace the ATTEND WORD and RETRIEVING rules with three rules ATTEND WORD, ENCODING WORD and RETRIEVING. The new rule ENCODING WORD mediates between ATTEND WORD and RETRIEVING: for our limited purposes, encoding a word form means taking it from the visual buffer and shuttling it to the imaginal buffer.

(8) Rule 5 (ENCODING WORD):

conditions				actions		
=goal>	STATE:	encoding	⟹	=goal>	STATE:	retrieving
=visual>	VALUE:	=val		+imaginal>	ISA:	word
					FORM:	=val

The ATTEND WORD and RETRIEVING rules are minimally revised: the output goal state for the ATTEND WORD rule is now ENCODING (rather than RETRIEVING), and the RETRIEVING rule looks up the string of characters in the imaginal buffer now (rather than the visual buffer).

(9) Rule 1 (ATTEND WORD; revised):

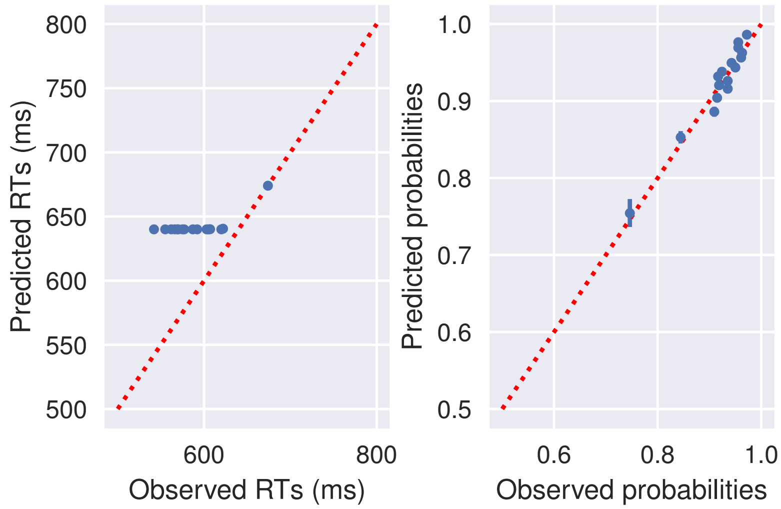
conditions				actions		
=goal>	STATE:	attend		=goal>	STATE:	encoding
=visual-location>	ISA:	-location	⟹	+visual>	CMD:	move-attention
					SCREEN-POS:	=visual-location
?visual>	STATE:	free				

(10) Rule 2 (RETRIEVING; revised):

conditions				actions		
=goal>	STATE:	retrieving	⟹	=goal>	STATE:	retrieval-done
=imaginal>	VALUE:	=val		+retrieval>	ISA:	word
					FORM:	=val

All these modifications are symbolic / discrete / qualitative. We are nonetheless able to fit the new model to the same data and quantitatively compare its performance with Model 1 (the no-imaginal-buffer model). As the left plot in Figure 4 shows, Model 2 has a very poor fit to the latency data. The encoding step adds 200 ms to every lexical decision, since 200 ms is the default ACT-R delay for chunk-encoding into the imaginal buffer. Consequently, the predicted latencies for 15 out of the 16 word-frequency bands are greatly overestimated (above the diagonal line). Model 2 cannot run faster than about 640 ms, and this is too high to fit high-frequency words, which take about 100 ms less than that.

Figure 4: Model 2: estimated and observed RTs and probabilities



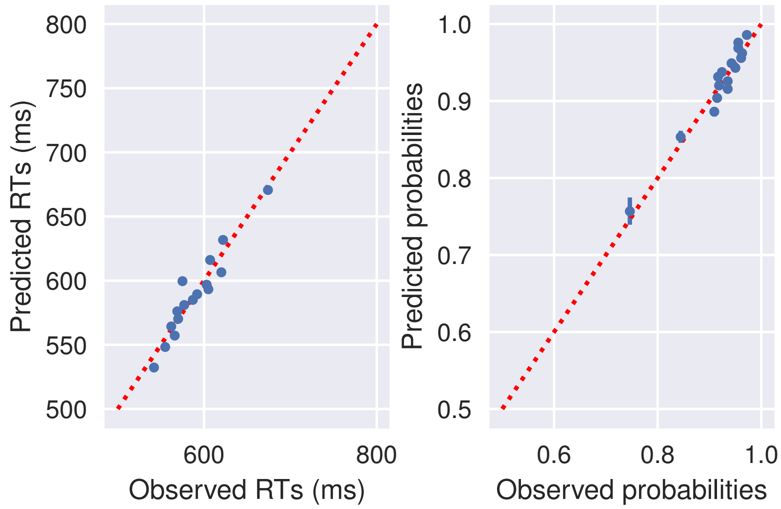
4.3 Model 3: imaginal buffer with 0 delay

Let's change a quantitative feature of Model 2 and set the imaginal delay to 0 ms instead. It is reasonable to assume that various default values for ACT-R subsymbolic parameters should be changed when modeling linguistic phenomena: natural language comprehension involves fast incremental construction of rich hierarchical representations, and this richness significantly exceeds the complexity of representations needed for other high-level cognitive processes modeled in ACT-R (e.g., arithmetic). This change is sufficient to obtain a very good fit to latencies for all 16 word-frequency bands, as shown in Figure 5.

5 Conclusion

We have presented a framework that integrates generative theories, the ACT-R cognitive architecture and Bayesian models. This framework enables us to do quantitative comparison for qualitative theories: we can implement different competence + processing models in ACT-R, and then embed these alternative ACT-R models in a Bayesian model. We can then estimate their subsymbolic parameters, and quantitatively compare these different models / theories. Consequently, we have a formally explicit way to connect competence-

Figure 5: Model 3: estimated and observed RTs and probabilities



level theories to experimental data via explicit processing models, and a formally explicit and systematic way to quantitatively compare these theories.

In this paper, we have done only informal quantitative comparisons based on posterior predictions, but systematic model comparison via WAIC values (as in Brasoveanu and Dotlačil 2018, for example) or Bayes factors is also possible. The framework can be used to model various other language-related processes, including eye fixation durations in eye-tracking-while-reading tasks (see Dotlačil 2018), and latencies in self-paced reading tasks targeting a variety of syntactic and semantic phenomena (Brasoveanu and Dotlačil 2018, Brasoveanu and Dotlačil 2019).

REFERENCES

Anderson, John R. (1982). "Acquisition of cognitive skill". In: *Psychological review* 89.4, p. 369.

Anderson, John R., Jon M. Fincham, et al. (1999). "Practice and retention: A unifying analysis". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25.5, pp. 1120–1136.

Anderson, John R. and Christian Lebiere (1998). *The Atomic Components of Thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brasoveanu, Adrian and Jakub Dotlačil (2018). "An extensible framework for mechanistic processing models: From representational linguistic theories to quantitative model comparison". In: *Proceedings of the 2018 International Conference on Cognitive Modelling*.

— (2019). *Formal Linguistics and Cognitive Architecture*. Language, Cognition, and Mind (LCAM) Series. Under review. The current version of the *pyactr* library (Python3 ACT-R) is available here: https://github.com/jakdot/pyactr; the frozen version used in the book, together with the code for all the models, is available here: https://github.com/abrsvn/pyactr-book. Dordrecht: Springer.

Dotlačil, Jakub (2018). "Building an ACT-R Reader for Eye-Tracking Corpus Data". In: *Topics in cognitive science* 10.1, pp. 144–160.

Forster, Kenneth I (1990). "Lexical processing". In: *Language: An invitation to cognitive science*. Ed. by Daniel Osherson and Howard Lasnik. Cambridge, MA: MIT Press, pp. 95–131.

Hart, Betty and Todd R Risley (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Paul H Brookes Publishing.

Howes, Davis H and Richard L Solomon (1951). "Visual duration threshold as a function of word-probability." In: *Journal of experimental psychology* 41.6, p. 401.

Lewis, Richard and Shravan Vasishth (2005). "An activation-based model of sentence processing as skilled memory retrieval". In: *Cognitive Science* 29, pp. 1–45.

Logan, Gordon D (1990). "Repetition priming and automaticity: Common underlying mechanisms?" In: *Cognitive Psychology* 22.1, pp. 1–35.

Monsell, Stephen (1991). "The nature and locus of word frequency effects in reading". In: *Basic processes in reading: Visual word recognition*. Ed. by D. Besner and G. W. Humphreys. Hillsdale, NJ: Erlbaum, pp. 148–197.

Murray, Wayne S and Kenneth I Forster (2004). "Serial mechanisms in lexical access: the rank hypothesis." In: *Psychological Review* 111.3, p. 721.

Whaley, Charles P (1978). "Word-nonword classification time". In: *Journal of Verbal Learning and Verbal Behavior* 17.2, pp. 143–154.

On the Paradox of Changting Hakka Tone Sandhi

LUHUA CHAO

*Tongji University & University of California, Berkeley**

1 Introduction

The subject of this paper is Changting, a dialect of Hakka (Kejia), one of the seven major Chinese dialects, mainly spoken in Changting county, located in the mountain areas of western Fujian province. This dialect possesses an inventory of five citation tones. Though no new tones are created, a large number of sequences, 15 out of 25 (=5^2) in disyllabic tonal combinations, henceforth BIPARTITE, and 108 out of 125 (=5^3) trisyllabic combinations, henceforth TRIPARTITE, are involved in tone sandhi, and the sandhi patterns are both various and complex, affecting either the first or the second tone or both tones of a T_1+T_2 sequence (Li 1965, Luo 1982, 2002, 2012, Rao 1986, 1987, Hsu 1995, Chen 2004, Chen et al. 2008). Previous studies have tried different approaches in order to explain its sandhi patterns, including generative approaches and classic Optimality Theory (OT; Prince & Smolensky 2004[1993], Kager 1999, McCarthy 2002, 2008) and Optimality Theory with Candidate Chains (OTCC; McCarthy 2007), but all turned out to be inefficacious (Yan 2016). Due to this fact, Chen et al. 2008 named it as THE PARADOX OF HAKKA TONE SANDHI and invited researchers to 'use the data as a testing ground for sharpening, extending, and perhaps radically reconceptualizing linguistic theory (4)'.

The paradox of Changting Hakka tone sandhi consists mainly of two parts. First, of the 108 out of 125 tritonal sequences undergoing sandhi, only 72 are explainable by the general sandhi rules, while the other 36 remaining unexplained. Second, the directionality of tone sandhi rule application seems random, not predictable by any conceivable factor. The goal of this paper is to provide a solution to the first part of the paradox, paving the way for tackling the second part, ultimately arriving at a resolution of the problem.

*Sponsored by China Scholarship Council (CSC).

The paper is organized as follows. In §2, I introduce the tone system of Changting. §3 presents the 15 out of 25 bipartite tonal alternations, accounted for by the previous tone sandhi (TS) rules which I refer to as TS1. I then show that TS1 accounts for 72 'explainable' tritonal sandhi cases. Instances of the still unexplained cases are given in §4, for which TS2 (a second type of tone sandhi differing from TS1) is introduced. TS2 can account for the other 36 tritonal sandhi cases. The two types of tone sandhi are further discussed in §5. Finally, based on the application of TS2, §6 reveals two features in the tone sandhi of the Changting dialect which are crucial to tackle the sandhi directionality, leading to the final resolution of the Hakka tone sandhi paradox.

2 The Changting tone system

Changting Hakka dialect has a five-tone system. Though previous investigations into the five citation tones vary, they can clearly be identified as /M, R, F, H, L/,[1] with M (=[L $_{[H]}$]) a H-registered low level tone, L (=[L $_{[L]}$]) a L-registered low level tone, R (=[LH $_{[L]}$]) a L-registered rising contour, F (=[HL $_{[H]}$]) a H-registered falling contour, and H (=[H $_{[H]}$]) a H-registered high level tone. Given that M and L differ only in register and share similar sandhi behaviors, they are grouped together as Subsystem I, and R, F, and H as Subsystem II.

3 Tone Sandhi 1 (TS1)

3.1 TS1 alternations

As schematized in Table 1 and exemplified in Table 2, of the 25 bipartite tonal combinations, 15 are involved in tone sandhi. Of these, 13 go through regressive modification of the first tone, one (R+L) goes through progressive modification of the second tone, and another sandhi process (F+L) affects both tones. This type of sandhi, here designated as TS1, is recognized in the above-cited literature.

[1]M=33, R=24~35, F=42~43, H=55~54, L=21~11, with Chao (1930) pitch integers of 1 (low) to 5 (high) assigned to the five tones. For H and L, Li (1965), though described in his own documents as 54 and 21, pointed out that these two tones are approaching level tones in shapes and should be recorded strictly as 554 and 221, rendering the differences between 54 and 55, 21 and 11 mere phonetic ones.

	M	R	F	H	L
M		L-R			L-L
R	H-M				R-F
F	R-M	L-R	M-F	L-H; F-M	R-F
H	F-M	F-R			F-L
L	M-M		M-F	M-H	

Table 1

TS1 of the Changting dialect (Chen et al. 2008)

	Base form	Sandhi form	Example	Gloss
Subsystem I+T_2	MR	LR	*fan.shu*	‘sweet potato’
	ML	LL	*zhong.xue*	‘middle school’
	LM	MM	*bai.ma*	‘white horse’
	LF	MF	*mao.xian*	‘to take risk’
	LH	MH	*jiu.huo*	‘old merchandise’
Subsystem II+T_2	RM	HM	*tao.hua*	‘peach blossom’
	RL	RF	*bei.ji*	‘north pole’
	FM	RM	*huo.che*	‘train’
	FR	LR	*wu.shi*	‘noon time’
	FF	MF	*xuan.ju*	‘to elect’
	FH	LH	*jian.jia*	‘to cut price’
		FM	*hao.xiao*	‘funny, laughable’
	FL	RF	*li.mao*	‘courtesy’
	HM	FM	*song.shu*	‘to give book’
	HR	FR	*xi.jie*	‘details’
	HL	FL	*ban.ye*	‘midnight’

Table 2

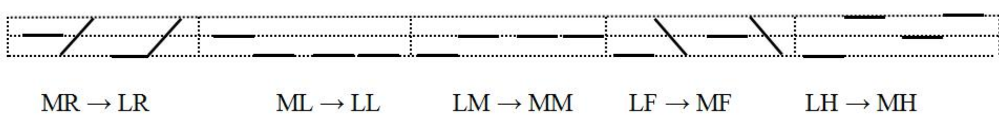
TS1 examples (Chen et al. 2008)

In order to account for TS1 I now address ‘Subsystem I+T_2’ and ‘Subsystem II+T_2’ in turn.

‘Subsystem I+T_2’ consists of combinations where the first tone is either L or M. As seen in Table 1, L and M undergo REGISTER ASSIMILATION and merge in complementary environments. M becomes L before L-registered L and R, and L becomes M before H-registered M, F, and H, as illustrated in 1.

(1) Register Assimilation (‘Subsystem I+T_2’ sandhi): M, L $\rightarrow$ L $_{[[L]]}$ / _ R, L $_{[L\text{-Register}]}$

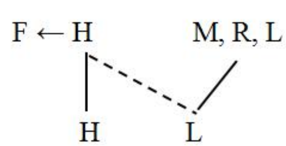
$\rightarrow$ M $_{[[H]]}$ / _ M, F, H $_{[H\text{-Register}]}$



Note: The three horizontal lines represent L, M, H, starting from the bottom line, and the vertical lines divide the different bitonal sandhi.

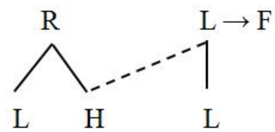
‘Subsystem II+T_2’ refers to bitonal combinations where the first tone is H, R, or F. When the first tone is H, it undergoes right to left (anticipatory) TONE ASSIMILATION from the second tone. The common L feature of M, R, and L spreads leftwards onto the tonal node of H, changing the H into a HL falling contour, as schematized in 2.

(2) Right to Left Tone Assimilation (H+T_2 sandhi): H $_{[H]}$ $\rightarrow$ F $_{[HL]}$ / _ M, R, L $_{[L\text{-Feature}]}$



When the first tone is R, RM and RL undergo sandhi. The change of RL to RF is a process of left to right (perseverative) tone assimilation. The H feature of R spreads onto the tonal node of L, changing it into HL, as in 3.

(3) Left to Right Tone Assimilation (RL to RF sandhi): $L_{[L]} \rightarrow F_{[HL]} / R_{[H\text{-}Feature]}$ _

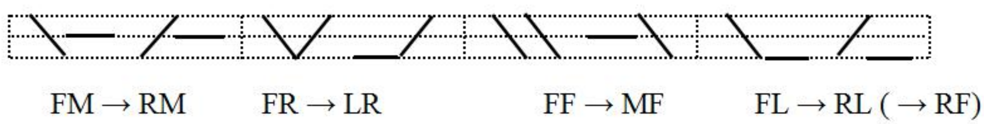


F+T_2 undergoes different sandhi according to whether the triggering tone is a level tone or a contour. As indicated in 4, when the second tone is a low level tone, that is M or L, F goes through CONTOUR METATHESIS, changing into R.[2] While when the second tone is a contour, that is R or F, F+T_2 undergoes CONTOUR LEVELING, F changing into L or M and register assimilating at the same time.[3]

(4) a. Contour Metathesis: $F_{[HL]} \rightarrow R_{[LH]}$ / _ M, $L_{[Level]}$

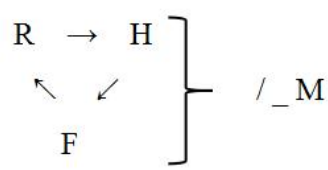
b. Contour Leveling: $F_{[HL]} \rightarrow L_{[L]}$ /_ R

$\rightarrow M_{[L]}$ / _ $F_{[Contour]}$



Before M, R changes into H, H changes into F and F changes into R, constructing a TONE CIRCLE, as in 5.

(5) A tone circle in the tonal environment of M:



[2]The bidirectional sandhi FL to RF is divided into two steps, the first step being FL to RL and the second step RL to RF, identical to the original progressive sandhi.

[3]FH sandhi is special in two aspects. First, FH to LH sandhi is not consistent with the generalizations of the other F+T_2 sandhi. If it were, it would undergo contour metathesis before level tone H and changes into RH. Second, different from the other 14 cases, FH has a second output FM and the sandhi of FH to FM is structure-sensitive, that is, when the tritonal sequence has a left-branching structure of [TF]H, the sandhi of FH to FM takes place at the juncture point. Otherwise, FH undergoes regressive sandhi and turns into LH.

Given the similarities between M and L, it is expected that before L, another tone circle is formed. However, instead of undergoing regressive sandhi and turning R into H, RL undergoes progressive sandhi, L changing into F, which blocks the potential formation of another tone circle, as illustrated in 6.

(6) A 'two-thirds' tone circle in the tonal environment of L:

$$\left.\begin{array}{ccc} \text{R} & & \text{H} \\ \nwarrow & & \swarrow \\ & \text{F} & \end{array}\right\} \ / \_\ \text{L}$$

3.2 Rule application in tripartite sandhi cases

Of the 108 (out of 125) tripartite tonal sequences that undergo sandhi, 72 are explainable by TS1 rule application, as shown, for example, in cases 7-10.

(7) MHM → MFM

E.g. *[cu.bu] shan* 'denim shirt'

chou [zhuang.ding] 'conscript (verb)'

$\underline{\text{MH}}\text{M}$ (⇨)	$\text{M}\underline{\text{HM}}$ (⇦)
n/a	|
$\text{M}\underline{\text{HM}}$	$\underline{\text{MF}}\text{M}$
|	n/a
MFM	MFM

(8) FLL → RFL

E.g. *[tao.lun] hui* 'seminar'

kao [da.xue] 'take college entrance exams'

$\underline{\text{FL}}\text{L}$ (⇨)	$\text{F}\underline{\text{LL}}$ (⇦)
| |	n/a
$\text{R}\underline{\text{FL}}$	$\underline{\text{FL}}\text{L}$
| |	| |
*RRF	RFL

Note: $\underline{\text{XY}}$ = the tonal substring being scanned or operated on.

⇨ , ⇦ left to right, right to left

The vertical line connects the target input (top) and the corresponding output (bottom).

n/a = not applicable * = unattested

(9) LFR → MLR

E.g. *[dai.biao] tuan* ‘delegation’

da [sheng.cheng] ‘big provincial capital’

LFR (⇨)	LFR (⇦)
|	|
MFR	LLR
|	n/a
MLR	*LLR

(10) LHR → MFR

E.g. *er.si.qi* ‘two, four, seven’

jiu [suan.pan] ‘old abacus’

LHR (⇨)	LHR (⇦)
|	|
MHR	LFR
|	|
MFR	MFR

In 7, only one rule, HM→FM, is applicable, and whether applied left to right or right to left, the same sandhi output is produced. In 8, FLL undergoes right to left sandhi and produces RFL as the attested sandhi output by applying the rule of FL to RF at the second step of the derivation. In 9, LFR undergoes two steps of sandhi in left to right direction and produces MLR as the attested sandhi output by applying the rules of LF to MF and FR to LR. In 10, though different rules are applied in two directions, they happen to produce the same sandhi output.

In similar ways, all of the 72 out of 108 tritonal sandhi cases are explainable in either left to right, right to left direction, or both directions, by applying TS1 rules at the first or second step or both steps of the derivation.[4]

4 Tone Sandhi 2 (TS2)

4.1 TS2 alternations

While two thirds of the tritonal sandhi can be accounted for by TS1, there are 36 sequences that remain unexplained. Considered exceptions in early studies, two of these are exemplified below.

[4]The tritonal sandhi data of Luo (2002) is applied. Luo is a native speaker of Changting, and her investigation into this dialect, primarily tone sandhi, has a long history, dating back to 1982.

(11) RFM → HRM ?

E.g. *[chu.ben] shu* ‘publish a book’

Huang [jie.zhang] ‘brother-in-law surnamed ‘Huang’’

RFM (⇨)	RFM (⇦)
n/a	|
RFM	RRM
|	n/a
*RRM	*RRM

(12) FFM → MRM / RRM ?

E.g. *[bao.xian] xiang* ‘safety box’

si [huo.shan] ‘extinct volcano’

FFM (⇨)	FFM (⇦)
|	|
MFM	FRM
|	|
MRM	*LRM

In 11, applied in either direction, TS1 yields RRM, which is unattested. In 12, FFM has two attested sandhi outputs. While MRM is obtained by a left to right derivation, how can RRM be derived?

It turns out that there is a generalization underlying the 36 exceptional cases. If the first tone is followed by a tone sandhied by TS1, a different set of outputs (TS2) is obtained. Recall the TS1 tone circle by which R changes into H, H changes into F and F changes into R. This tone circle is what TS2 abides by when the middle tone is sandhied by TS1, as shown in Table 3 below.

	M	R	F	H	L
R	H-M	H-R	H-F	H-H	R-F $_{H-L}$
F	R-M	R-R	R-F		
H	F-M	F-R	F-F	F-H	F-L

Table 3

TS2 of the Changting dialect

As indicated in Table 3 and recapitulated in Table 4, TS2 is further divided into four classes according to how it relates to its TS1 counterparts. In Class I, the sandhi of RM, FM, HM, HR, and HL (indicated by the unshaded cells) takes place in accord with the tone circle, hence their TS1 and TS2 coincide. Class II is the sandhi of RR, RF, RH, HF, and HH (indicated by the cells shaded in light gray). For this class, only TS2 exists. Class III is the sandhi of FR and FF (indicated by the cells shaded in dark gray). Their TS1 and TS2 both exist, but differ. Class IV is the sandhi of RL. TS2 of RL is restricted in that only in the cases where TS1 of RL to RF is not allowed to take place, TS2 of RL to HL gets a chance to be realized (indicated by RF footnoted by HL). FH and FL have no TS2 rules. In a derived environment (_T'), their TS1 rules apply.

Class	Sandhi rules	Relations with TS1	Rules applied in _T' context
I	RM - HM; FM - RM; HM - FM; HR - FR; HL - FL	Identical	TS2
II	RR - HR; RF - HF; RH - HH; HF - FF; HH - FH	No TS1 counterpart	TS2
III	FR - RR; FF - RF	Different	TS2
IV	RL - HL	Different	TS1 prior to TS2

Table 4

The four classes of TS2

4.2 Rule application in tripartite sandhi cases

Putting TS2 into the derivational routes of 11 and 12, the unexplained cases of RFM and FFM are explained. In 11, at the second step of the right to left derivation, TS2 of RR to HR (Class II) takes place, and as a result, the attested sandhi output of HRM is produced. In 12, TS2 of FR to RR (Class III) takes place, bringing the second attested output to the FFM sandhi. Since the second step has an input derived from TS1, TS2 rule gets a chance to be applied.

The following cases 13-15 demonstrate how the TS2 rules of Class I and Class IV apply.

(13) RLM → HMM

E.g. *[shao.yao] hua* ‘peony’

huang [han.shan] ‘yellow undershirt’

RLM (⇨)	RLM (⇦)
|	|
RFM	RMM
|	|
*RRM	HMM

Note: XY = the tonal substring subject to TS2 rule application.

In 13, the attested sandhi output HMM is produced by applying the TS2 rule of RM to HM (Class I) triggered by the sandhied M at the second step of the right to left derivation.

(14) RML → RFL

E.g. *[lian.huan] hui* ‘festival’

cai [shan.xiu] ‘cut sleeves’

RML (⇨)	RML (⇦)
|	|
HML	RLL
|	|
*HLL	RFL

(15) RFR → HLR

E.g. *[pi.jiu] ping* ‘beer bottle’

chen [jing.lin] (a person’s name)

RFR (⇨)	RFR (⇦)	RFR (⇦)
n/a	|	|
RFR	RLR	RLR
|	|	|
*RLR	HLR	*RFR (identical to the input)

In 14, RL undergoes TS1 regardless of the fact that L is a sandhied tone. While in 15, applying TS1 rule would produce a potential output RFR, which is identical to the original tritonal string. In order to avoid this, TS2 of RL to HL takes place instead.

As TS2 adding into the sandhi catalog, no case in Changting Hakka tripartite tone sandhi is rendered unexplained any more. The first part of the paradox is, therefore, solved.

5 TS1 versus TS2 in tritonal sequences

In this section, I recapitulate the characteristics of TS2 by contrasting it with TS1.

In a bipartite tonal combination of T_1T_2, two different contexts are distinguished. One is an underlying, unchanged tone, that is a tone constituting the original tritonal string, represented by T, while the other is a derived tone. A derived tone, represented by T', is a tone sandhied by TS1.

Before both T and T', the tones in Subsystem I, that is L and M, as well as F in FH and FL undergo TS1, due to the fact that these tones have no TS2 rules. In contrast, Subsystem II tones, that is R, F, and H, undergo different sandhi according to different tonal contexts, going through TS1 when preceding T while TS2 when preceding T'.

As seen, TS1 and TS2 are both sensitive to neighboring tones, but in different ways. To be specific, TS1 cares about the specific identity of the adjacent tone and takes place by interacting with its different tonal components. On the other hand, TS2 is triggered by a derived tone, with no regard to the specific identity of the derived tone, and takes place in accord with a tone circle.

Speaking of visibility, TS1 is more immediately apparent or overt in that it is observable from a simpler bipartite input. Compared to TS1, TS2 is more subtle or covert. It does not show itself unless the derivation of a tritonal sandhi reaches its second step and the condition is met.

Of the 108 out of 125 tritonal sequences undergoing sandhi, 72 are explainable by applying TS1 rules only. On the other hand, the remaining 36 cases rely on both TS1 and TS2 interacting with each other, cooperatively and competitively.

The different characteristics of TS1 and TS2 are listed in Table 5 below.

	Aspects	TS1	TS2
①	TS context	Subsystem I+T/T'; Subsystem II+T; F+H', F+L'	Subsystem II+T'

②	TS manner	Relevant to the specific identity of the adjacent tone	In accord with the tone circle
③	TS visibility	Immediately apparent or overt	Subtle or covert
④	Involvement in tripartite sandhi cases	In all 108 out of 125 cases; 72 with TS1 involvement solely	36

Table 5

TS1 versus TS2

6 Conclusion

In the preceding sections we have seen that TS2 plays a major role in tone sandhi of Changting Hakka dialect. In this brief conclusion I will show that there exist two inner features in the tone sandhi of this dialect, without which the application of TS2 rules would be impossible.

Recall the FFM case in 12. At the second step of the right to left sandhi, before derived R, TS2 of FR to RR takes place. However, in 16 below, since R is not derived by the previous step, but an underlying tone, FR undergoes TS1 and changes into LR.

(16) FRL→ LRF

E.g. *[wu.ping] yao* ‘five bottles of pills’

gu [wen.zi] ‘ancient characters’

FRL (⇨) FRL (⇦)

| |

LRL FRF

| |

LRF LRF

This tells us that a tone sandhi has to be able to get access to the derivational history of the adjacent tone in order to decide which type of sandhi it undergoes. This is what Kenstowicz and Kisseberth (1977) refer to as LOOK-BACK GLOBALITY.

Then recall the RML and RFR case in 14 and 15. In 14, TS1 of RL to RF takes place even though L is a derived tone. However, in 15, if TS1 rule were applied, the output RFR, which is exactly the same as the tritonal input, would be produced. To avoid this, RL undergoes TS2 and changes into HL. This is LOOK-AHEAD GLOBALITY. Different from look-back globality, look-ahead globality requires a tone sandhi to be capable of foreseeing what is going to happen next, as well as the potential sandhi output, before it chooses the appropriate type of sandhi to go through.

The Changting tone sandhi exhibits both the features of look-back globality and look-ahead globality, which have different functions. Look-back globality enables TS2 to take place, which is responsible for the explanation of the 36 unexplained tritonal cases. Look-ahead globality, on the other hand, crucially regulates the sandhi directionality, ultimately leading to a resolution of the Hakka tone sandhi paradox.

REFERENCES

CHAO, YUEN-REN. 1930. A system of tone letters. *La Maître phonétique* 45.24-27.

CHEN, MATTHEW. 2004. Changting Hakka tone sandhi: analytical challenges. *Language and Linguistics* 5.799-802.

CHEN, MATTHEW; LIAN-HEE WEE; and XIU-HONG YAN. 2008. The paradox of Hakka tone sandhi. In Hsiao, and Yuchau, E. (eds.) *Interfaces in Chinese phonology*. Taiwan: Institute of Linguistics, Academia Sinica. 1-45.

HSU, HUI-CHUAN. 1995. Trisyllabic tone sandhi in the Changting Hakka dialect. *Journal of Chinese Linguistics* 23.42-86.

KAGER, RENÉ. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

KENSTOWICZ, MICHAEL, AND CHARLES KISSEBERTH. 1997. *Topics in phonological theory*. Academic Press.

LI, RU-LONG. 1965. Disyllabic and trisyllabic tone sandhi in Changting dialect. *Journal of Xiamen University (Arts and Social Sciences)* 2.57-61.

LUO, MEI-ZHEN. 1982. Tone sandhi of Hakka dialect in Changting, Fujian. *Studies in Language and Linguistics* 2.188-97.

LUO, MEI-ZHEN. 2002. Tone sandhi of Hakka dialect in Changting, Fujian: a supplement. *Studies in Language and Linguistics* 4.91-97.

LUO, MEI-ZHEN. 2012. Some characteristics of Hakka dialect in Changting, Fujian. *Journal of Longyan University (Arts and Social Sciences)* 30.1-5.

MCCARTHY, JOHN J. 2002. *A thematic guide to optimality theory*. Cambridge: Cambridge University Press.

MCCARTHY, JOHN J. 2007. *Hidden generalizations: phonological opacity in optimality theory*. London: Equinox Publishing.

MCCARTHY, JOHN J. 2008. *Doing optimality theory: applying theory to data*. Oxford: Blackwell.

PRINCE, ALAN, AND PAUL SMOLENSKY. 2004[1993]. *Optimality theory: constraint interaction in generative grammar*. Oxford: Blackwell.

RAO, CHANG-RONG. 1986. A brief account on the phonetics of Hakka dialect in Changting, Fujian. *Journal of Longyan University (Arts & Social Sciences)* 2.30-42.

RAO, CHANG-RONG. 1987. Tone sandhi of Hakka dialect in Changting, Fujian. *Studies of the Chinese Language* 3.

YAN, XIAO-BIN. 2016. OTCC and the study of Changting Hakka tone sandhi. *Journal of Shandong Agricultural University (Social Science Edition)* 18.103-8.

The pragmaticalization and synchronic variations of Japanese adverb *jitsuwa* and English adverb *actually*

KIYONO FUJINAGA
University at Buffalo, SUNY

1 Introduction

If you look up *jitsuwa* in a Japanese-English dictionary and *actually* in an English-Japanese dictionary, the two adverbs are often treated as translation equivalents. As L2 learners, mastering these adverbs is tricky because they contribute to the speaker's attitude towards the presented proposition or towards the interlocutors, rather than the semantic aspects of the proposition. On top of this, the learners have to pay attention to where these adverbs are used in a clause because the position cues the speaker's communicative intention. This paper argues that although the Japanese adverb *jitsuwa* and the English adverb *actually* are often translated equivalently, each adverb has synchronically variant meanings at different pragmaticalization stages. This is why pragmatic failures occur when L2 learners assume only one meaning/function of these adverbs. Japanese learners whose first language is English are observed to make the pragmatic errors as in the example (1), where the choice of *jituwa* by the L2 speaker seems to originate in the English direct translation of *actually*: the L1 speaker asked the L2 speaker a question in line 23 as to what kind of conversations the L2 speaker used to have with his/her host family in Japan. In line 24, the L2 speaker explained why they would always end up talking a lot, rather than answering the question about the contents of the conversation. In line 25, the L1 speaker brought up the question again, this time more explicitly, asking what kind of conversations the L2 used to have with his/her host family. The L2 speaker started answering the question using *jitsuwa*, as if s/he is elaborating on his/her earlier answer in line 24.

(1) free style conversation between Speaker1 (L1 Japanese) and Speaker2 (L1 English; L2 Japanese)

23: Seaker1: asou desu ka. dou desu ka it-tara ano S2 ga it-tara ojiisan to obaasan

wa donna koto o hanashi-te kure-masu ka?
'Is that so? How about, well, when you visit Mr & Ms X, what kind of things did they talk about with you?'

24: Speaker2: anoo inaka ni sun-de i anoo imasu kara uu mmm anmari ano takusan koto ano dekiru koto amari ari-masen kedo ano takusan mmm tabe-tari ironna hanashi ga at-tari ano watashi no itoko no kodomo to mm ano asobi mm aso-n-dari a demo tanoshi
'Well, (they) live in the countryside, so uhm there were not that many things we could do, but uhm we would eat a lot, and there were conversations... uhm I would play with my cousin's children... hmm it was fun..'

25: Speaker1: aa sou desu ka. ee ano donna koto o isshoni hanashi-masu ka. ojiisan to obaasan to *'Alright. Well, what kind of conversations did you have with Mr & Ms X?'*

26: Speaker2: ee **jitsuwa** watashi no ano koyuto no seikatsu nitsuite, anoo soremo tabun watashi no amerika no kazoku nitsuite. anoo sore igai ni nn maa tabun ano nihongo o aa ima nihon no nyuuzu toka ano sekai kankei toka tokidoki no
'Well, what we really talked about was my life in Kyoto and uhm also maybe about my family in the US. Other than that, well, uhm Japan oh current Japanese news and global topics.'

The way the L2 speaker uses *jitsuwa* in line 26 of example (1) sounds bizarre because the speaker starts the utterance as if it is a confession. However, what follows *jitsuwa* is not really surprising news. This is a case of pragmatic failure, where a particular expression in the target language is perceived differently than the speaker's intention (Thomas, 1987). Similarly, learners of English are prone to pragmatic failures where the choice of *actually* seems to originate in direct translation from the Japanese adverb *jitsuwa*. As a result, such mis-use of *actually* sounds impolite. Take the following example, where an advanced L2 English speaker (L1: Japanese) is responding to a question on a standardized test:

(2) Question: Do you think smoking should be banned at restaurants?

Response:

> Hello, I I don't agree with idea but to ban to ban smoking all all the restaurant at all restaurants in the country. **Actually actually** I don't like smoking. I don't want I don't want to get smoking just I don't want to get smoke I don't want smoke emitted that emit emit in smoking. But the smoking is smoking is the lifeline, so-called for for the people who is smoking.

(Japanese intermediate English learner, ICNALE corpus)

In example (2), the speaker uses the adverb *actually* as if s/he is opposing a presupposition that is s/he smokes. One of the functions of *actually* is indeed showing contradiction, as we will see in section 2, and this opposition would be considered abrupt. The following study examines the usage of the Japanese adverb *jitsuwa* and the English adverb *actually* through primarily qualitative methods. In section 2, I will review previous studies regarding pragmaticalization, discourse markers and the meanings/functions of *actually*. In section 3, I will introduce my study, looking at Japanese and English CALLHOME corpuses (Canavan & Zipperlen, 1996; Canavan et al., 1997) in order to find out the pragmatic functions of the Japanese adverb *jitsuwa* and the English adverb *actually*. In section 4, the Japanese adverb *jitsuwa* will be examined in a monologue context . In section 5, I will discuss the relevant findings of the study and section 6 concludes the paper.

2 Previous studies

2.1 Pragmaticalization and discourse markers

Pragmaticalization is the development of various particles with discourse functions (Traugott, 1995, p.5). It is also characterized as development towards less truth-conditional and referential meanings (Ajimer, 1996). Although defining pragmaticalization can be problematic because of its approximation with grammaticalization, some scholars view pragmaticalization as part of grammaticalization by incorporating the discourse function into the grammar (Diewald, 2011). The key notion of this study is that, like grammaticalization, pragmaticalization is motivated by polysemy. The development from referential to more abstract meanings does not occur abruptly. Rather, both grammaticalization and pragmaticalization are gradual processes where the old meanings/functions and newly emerged meanings/functions synchlonically co-exist.

Discourse markers make relations between sequentially dependent units of discourse such as *so*, *and*, *because*, and *you know* (Shiffrin, 1987). Discourse markers, as a subtype of discourse particles, are the products of grammaticalization and pragmaticalization (Traugott, 1995). They tend to develop from textual > inferential > interpersonal meanings. Because pragmaticalization is supported by the polysymous of meanings/functions, a discourse marker can present various meanings/functions. Interestingly, as we will see in the next section, the development of these meanings corresponds to where they occur in a clause.

2.1.1 Meanings and functions sensitive to discourse position

Beeching & Detges (2014) argue that the linguistic elements appearing at the left and the right peripheries show different functions in discourse. The linguistic elements appearing at the left periphery, which is the beginning of an utterance, have been studied under the notion of information structure. That is, they organize a discourse coherence. Shiffrin (1987), for example, has developed the study of discourse markers by looking exclusively at those appearing at the beginning of an utterance, managing local and global level of discourse structures.

Scholars investigating English discourse markers have also been noticing that those discourse markers which tend to appear at the right periphery of an utterance serve different functions than those at the left periphery; they serve an interpersonal function in discourse. Let us consider the example (3).

(3) Discourse marker English *then* depending on different positions in a sentence

(medial) We went to the cinema **then** to a restaurant.
(left) **Then** you'll be tired.

(right) You're coming with us **then**?

(Beeching & Detges, 2014: 7)

When *then* occurs in the sentence medial position, it serves as a connective between two clauses, indicating the sequentiality of events. When *then* occurs at the left periphery of the sentence, it serves as an inferential marker, connecting previous discourse to the current utterance. When *then* occurs sentence-finally, it invites the hearer's involvement to the proposition. This development of discourse marker from medial position > left periphery > right periphery has also been observed in Japanese.

(4) Discourse marker Japanese *kara* depending on different positions in a sentence

(a) abunai **kara** yame yo.
dangerous because stop SF
'Let's stop it because (it is) dangerous.'

(b) da**kara** yame yo.
COP.because stop SF
'(I'm telling you) stop it.'

(c) yame yo abunai **kara**.
stop SF dangerous because
'Let's stop it. Cuz (it is) dangerous.'

(Higashiizumi, 2015: 140)

When *kara* appears in the sentence medial position, it shows the causality between the two propositions. When *kara* appears sentence-initially with copula verb *da*, it implies the causality between the previous discourse and the following utterance and the hearer is expected to connect these two. Thus, it also serves as attention-getter, indicating that the hearer must pay attention to the following information. When *kara* appears sentence-finally, the speaker communicates an engagement with the hearer.

2.1.2 Meanings and functions of *actually*

Ajimer (1986) was the first to point out that the English adverb *actually* has different functions, depending on the position it occurs in a sentence. *Actually* in medial position is realized as actual truth, as in (5),or as an emphasizer having a reinforcing effect on the truth value of the clause, as in (6).

(5) actual truth

He **actually** said something totally different I'm just being polite to

(Ajimer, 1986: 120)

(6) *actually* as an emphasizer

called Wilfred Evans who's **actually** a very nice chap

(Ajimer, 1986: 122)

When *actually* occurs at sentence-initial position, it signals or cues to the listener how two utterances are related to each other. At this position, functions such as explanation, elaboration, self-repair, and contradiction are observed.

(7) *actually* as elaboration

I'm not a good communicator, but it doesn't mean I'm no:t still her frie:nd
(0.4) **actually** I'm a pretty shitty communicator

When *actually* occurs at sentence-final position, it expresses a personal opinion rather than a fact.

(8) *actually* expressing a personal opinion

It's a pity **actually**.

It also functions to invite intimacy and sharing, "I'm telling you this as a confidence."

(9) *actually* showing intimacy

I was determined to get married **actually**.

3 The study

While previous studies analyzed the English discourse marker *actually* depending on its position in the clause, there is no descriptive study of the Japanese adverb *jitsuwa* as a discourse marker. Also, from the learner's point of view, it is not clear what discourse meanings/functions of these adverbs are most frequently realized and to what extent the position of these adverbs really favors the interpretation of meaning. By using the same corpora for both English and Japanese, the present study will examine how speakers use these discourse markers in conversation.

3.1 Corpus

I used the CALLHOME corpus of North America and Japan. The corpus was first published in 1997. The participants were native speakers of each language, asked to speak with their friends or family members for 30 minutes over the phone. Because of the nature of the interlocutors, their conversations were casual. Both CALLHOME corpus of North America and that of Japan consist of the 120 unscripted conversational dialogs. I counted the occurrence of *actually* in the North America version and *jitsuwa* in the Japanese version. I collected 84 occurrences of *actually* in English and 10 occurences of *jitsuwa* in Japanese.

3.2 Coding

I coded the occurrences of *actually* and *jitsuwa* in terms of where they appeared in a given utterance and what pragmatic function they served in the utterance. Following Ajimer's (1986) study, the position in the utterance was determined by whether they are sentence initial, medial or final.

(10) coding *actually* for the position

(initial) yea:h right, that's on:e thing: **actually**
(medial) things are **actually** happening now
(final) hhh u:m, not that often **actually**

I have coded the occurrences by the following functions: explanation/elaboration, emphasis, self repair,and contradiction.

(11) coding *actually* for the pragmatic function

(self repair) ummm **actually** I'll I'll give you the phone number
(contradiction) hhh u:m, not that often **actually**
(explanation/elaboration) ihh they **actually** have a ne:w(0.6) pla:ce in Minneapolis called
(actual truth) at first I wasn't sure if it was actually him laying do:wn
(emphasizer) ihhh and I can like **actually** start working on my bike
(personal opinion) u::m, (0.4) I (.) doubt it actually

3.3 Results

The results of the data suggests the position where *actually* occurs in an utterance indeed favors the specific meanings/functions. Especially, when *actually* occurs sentence-initially, it is used as self-repair, contradiction, explanation and elaboration. Notice that the difference between these four functions overlap. The difference between self-repair and contradiction is whether the same speaker continues his/her turn with the utterance starting with *actually*. That is, when the speaker uses utterance-initial *actually* within his/her turn, it is self-repair and if it is after someone else's turn, it is a contradiction and it challenges the other person's presupposition. Elaboration and explanation can be used within the same speaker's turn. When they are used for someone else's turn, it is a collaborative utterance. What is shared by these functions of *actually* is that they all connect previous utterances to the current utterance– a unique information structural property at the right periphery. Similarly, when *actually* is used at utterance-medial position, it is interpreted as either actual truth or emphasizer. Again there is an overlap between these two functions, but they are qualitatively different from those functions found sentence-initially. *Actually* at utterance-final position was rarely used.

10 occurrences of *jitsuwa* were all exclusively used at utterance initial, signaling the following information is newsworthy.

(12) *jitsuwa* as a confession marker
jitsuwa iu to jibun no heya osa:
'To tell you the truth, I did something to my room'

In order to find out whether *jitsuwa* is only used as a confession marker in other genres than conversation, I will examine the functions/meanings of *jitsuwa* in monologue in the next section.

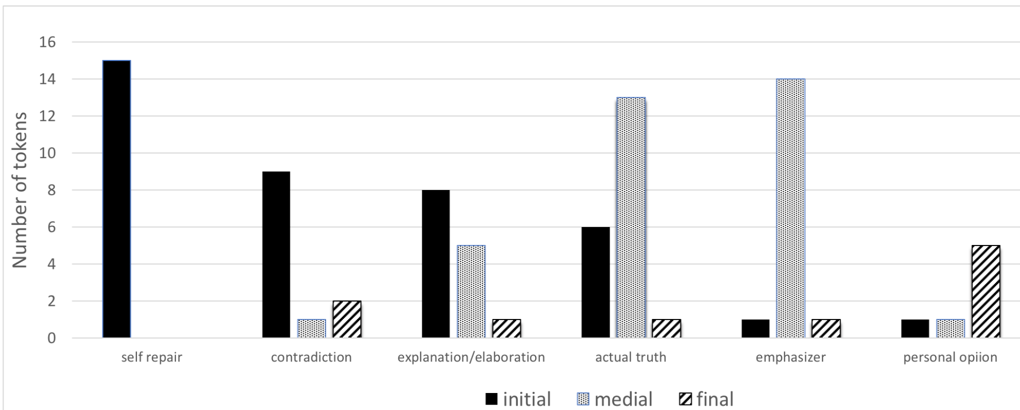


Figure 1: *actually* in Northern American CALLHOME corpus

4 *Jitsuwa* in monologue

In the Japanese CALLHOME corpus, which consists of dialogue, only 10 occurrences of j*itsuwa* were observed. Moreover, all the *jitsuwa* occurrences were used to indicate that the utterance was some kind of confession. In this section, I will provide a follow-up study to see whether *jitsuwa* occurs rarely in discourse and whether *jitsuwa* carries different pragmatic functions other than a confession.

4.1 Corpus

In this follow-up study, I used NIHONGO HNASHI-KOTOBA corpus 'Spoken Japanese Corpus'(CSJ) published by the National Institute for Japanese Language and Linguistics. The speech was collected from 1930 to 1970. For this study, I chose only monologue data, consisting of conference talk, lecture talk, and free talk. Among about 2000 occurrences of *jitsuwa*, I randomly selected 500. Unfortunately, the transcriptions are not available that inonation units are no clear.

4.2 Data analysis

Three types of pragmatic meanings/functions of *jitsuwa* were observed in the monologue data. First, just like *actually* is used as an emphasizer, *jitsuwa* can be used to highlight the significance of information. When used as an emphasizer, it does not present actual truth or surprising information.

(13) *jitsuwa* as an emphasizer

[...] kore **jitsuwa** hijouni anoo iroirona kou oshiro ga douiu huuni kou kuhuushi-te i-ta ka toiu koto o shiru koto ga dekiru genjou demooo taihen yoi ma iseki to itte yoi to omo i-masu souit-ta imi dewa[...]
'These are really well.. exceptional ruins today where we can learn how castles used to utilize the materials....'

Second, *jitsuwa* is used for providing new information in a monologue context. Related to *jitsuwa* as a confession marker in conversation. *jitsuwa* as new information providing marker in monologue is distinctively different from the confession marker found in conversation in the sense that the information is not entirely shocking. This is probably because the monologues in this corpus are very academic, such as lectures, and that these are qualitatively different from everyday relaxed conversation. However, when *jitsuwa* is used for providing new information, the speaker shifts to a rather casual speech style within a formal discourse genre, as seen in the use of the sentence final particle *ne*, as in (14).

(14) *jitsuwa* providing new information

[...] ano e eigo deni kagirazu nani-go demo desu ne joutatsu suru saikou no ichiban no houhou na n desu de watashi mo souyatte **jitsuwa** eigo to furansu-go o oboe-ta n desu **ne** [...]

'well not just English but for every other language, this is the best way to master a language. I learned English and French in this way. I did not use Japanese at all.'

Lastly, *jitsuwa* was found to be most frequently used to provide background information. Noticeably, the whole clause, where *jitsuwa* is used to provide background information, is low in transitivity. The particle *-ga* 'but', occurring in the same *jitsuwa* clause, is not used as contrastive, but rather for providing background[1], followed by the speaker's main claim.

(15) *jitsuwa* providing background

[...]kono kiki-kata wa **jitsuwa** eigo no rissun de wa nai toiu shiteki wa sudeni nasare-te i-masu **ga**, watashi wa hiaa toiu no tomo sukoshi chigau nodewanai ka to kanjite imasu[...].

'It has been already pointed out that this is not what English people would call "listen". I think it is not quite "hear" either.'

In monologue, we can see that *jitsuwa* is synchlonically more varied than in the conversation, where the only function observed was confession.

5 Discussion

From the CALLHOME corpus data, I established that the English adverb *actually* has been pragmaticalized and represents different pragmatic meanings at different positions in a clause. Most importantly, these meanings/functions co-existed, except functions found at the sentence-final position, such as softening and expressing personal opinion, which were rarely used. *Jitsuwa*, on the other hand, was used exclusively as a confession marker at the beginning of the utterance. This suggests that if the L2 Japanese learner (L1: English) believes that *jitsuwa* and *actually* are translation equivalents, they could fail in their intended meaning. Hence, this may lead to potential pragmatic failure because the L1 Japanese speakers will always understand utterance-initial position *jitsuwa* as a confession marker and expect a shocking story to follow the conversation. This accounts for the

[1] see for the detailed study in non-contrastive use of *kedo* 'but' in Nakayama & Nakayama (1997)

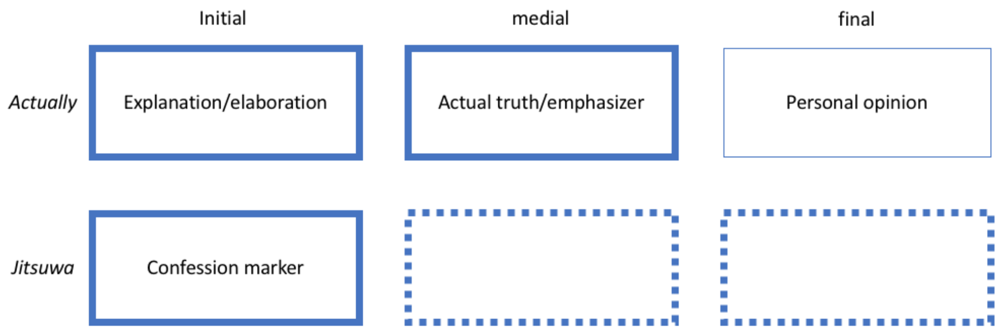


Figure 2: *actually* and *jitsuwa* in dialogue

L2 speaker's bizarre use of *jitsuwa* in the example (1). Similarly, when the L2 English learner (L1: Japanese) uses *actually*, meaning *jitsuwa*, they could be misunderstood as in example (1). As we have seen in section 3.3, utterance-initial *actually* can be understood as self-repair or contradiction of the interlocutor. Contradiction of the speaker him/herself is self-repair while contradiction of the speech interocular is a potential face-threatening act. For this reason, the L2 learner's use of *actually* may sound challenging to the English speaker.

Example (2) is a monologue , but there is also a conversational tone since it is a response to a given question. It is not clear from example (2) which of the following pragmatic meanings of *jitsuwa* would best fit the L2 speaker's intended meaning: confession, new information or background. Nevertheless, these three meanings overlap because of the polysemous relationship caused by the grammaticalization/pragmaticalization process. *Jitsuwa* can be divided into a noun *jitsu* 'truth' + a case particle *wa*, although these two segments are usually interpreted as one chunk. The *wa* particle itself seems to be the product of grammaticalization. In present day Japanese, the *wa* particle is used as a topic marker as well as a contrastive marker. Previous studies (De Wolf, 1987; Fujii-Ueno, 1987) have found that the *wa* particle developed as a contrastive marker and later became a more neutral topic marker. *Wa* in *jitsuwa* as confession marker seems to retain this old contrastive meaning. It contrasts the common ground between the interlocutors and the actuality. *Jitsuwa*, as a new information marker, seems to have lost its contrastive meaning, which the confession marker shows. *Jitsuwa* as a background marker seems to have gone far beyond loosing contrastive meaning– it does not even present a topic, but rather presents background information about a main topic.

6 Conclusion

This paper discussed that although the Japanese adverb *jitsuwa* and the English adverb *actually* are often translated equivalently, each adverb has different synchronically variant meanings at different positions in a clause. This is why pragmatic failures occur when learners assume only one meaning/function of these adverbs. For L2 learners of Japanese, special caution must be used when employing *jitsuwa* at utterance-initial position in conversation since this cues L1 speakers to receive shocking news. As for L2 learners of English, the use of *actually* at utterance-initial position is a potential face-threating act to the interlocutor. Although the study found various synchronic variations for the Japanese adverb *jitsuwa*, future studies are needed from diachlonical perspectives to confirm whether the development of the case particle *wa* plays a role, and whether the low transitivity of the *jitsuwa* clause is an on-going process. In addition to this analysis of the relation of meaning to the position of clauses for both usages of *actually* and *jitsuwa*, further investigation into differences between monologues and dialogues will also be needed.

REFERENCES

AJIMER, KARIN. 1996. I think -an English modal particle. *Modality in Germanic Languages: Historical and Comparative Perspectives*, ed. by Toril Swan and Olaf Jansen Westvik, 1-47. Berlin: Mouton de Gruyter.

BEECHING, KATE and DETGES, ULRICH. (eds.) 2014. *Discourse Functions at the Left and Right Periphery: Crosslinguistic investigations of language use and language change.* Leiden/Boston: Brill.

DE WOLF, CHARLES M. 1987. *Wa* in diachronic perspective. *Perspectives on Topicalization*. eds. by John Hinds; Shoichi Iwasaki and Senko K. Maynard, 265-290.

Amsterdam: John Benjamins Publishing Company.

DIEWALD, GABRIELE. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49. 365-390.

FUJII-UENO, NORIKO. 1987. *Wa* from synchronic and diachronic perspectives. *Perspectives on Topicalization*. eds. by John Hinds; Shoichi Iwasaki and Senko K. Maynard, 221-262. Amsterdam: John Benjamins Publishing Company.

HIGASHIIZUMI, YUKO. 2015. Periphery of utterances and (inter) subjectification in Modern Japanese. *New Directions in Grammaticalization Research.* eds. by Andrew D.M. Smith; Graeme Trousdale and Richard Waltereit, 135-154. Amsterdam: John Benjamins Publishing Company.

SCHIFFRIN, DEBORAH. 1987. *Discourse markers*. Cambridge: Cambridge University Press.

TRAUGOTT, ELIZABETH. C. 1995. The role of discourse markers in a theory of grammaticalization. Paper presented at the 12th International Conference on Historical Linguistics, Manchester.

TRAUGOTT, ELIZABETH. C. 2010. Revisiting subjectification and intersubjectification. *Subjectification, Intersubjectification and Grammaticalization*, ed. by Kristin Davidse; Lieven Vandelanotte and Hubert Cuyckens, 2970. Berlin: De Gruyter Mouton.

Subanon *mo-* cancels out volitionality: Evidence from paradigms and argument structure

BRYN HAUK
*University of Hawai'i at Mānoa**

1 Introduction

This paper addresses a descriptive problem in Subanon [ISO 639-3: suc], an underdescribed language spoken on Mindanao island in the Philippines. Specifically, I describe the Subanon prefix *mo-* by establishing its place in a paradigm, which I call non-volitional, with values for mood (irrealis, realis) and voice or focus (patient voice (PV), actor voice (AV), and goal voice (GV)), as shown in Table 1. I further relate this non-volitional paradigm to another paradigm, called volitional, following an analysis of the cognate prefix *ma-* in Tagalog by Himmelmann (2006). The Subanon non-volitional closely resembles what Himmelmann called "potentive" in Tagalog. For Himmelmann, potentive referred to forms that can be interpreted as denoting an accidental action, involuntary action or an ability ascribed to the most agent-like argument. The Subanon volitional paradigm is analogous to Himmelmann's dynamics in Tagalog.

Table 1: The Subanon non-volitional paradigm

	irrealis	realis
PV	**mo-**	mi-
AV	moko-	miko-
GV	ko-...-an	ki-...-an

However, the non-volitional paradigm alone does not resolve the descriptive problem in Subanon, because it does not fully predict the meaning and patterning of a predicate with

*My sincere thanks to Sharon Estioca for allowing me to contribute in this small way to her important work documenting her language and for the time she spent sharing her language with me. I also wish to thank William O'Grady, Robert Blust, Gary Holton, Lyle Campbell, my colleagues in the 2015-2016 field methods class at the University of Hawai'i at Mānoa (UHM), and the audience of at the 44th Annual Meeting of the Berkeley Linguistic Society for their helpful comments. My participation in this conference was made possible by support from the Graduate Student Organization at UHM (award number 18-03-02).

this type of affixation, the main desideratum for a linguistic paradigm. Subanon predicates bearing this affixation can denote any of the following: potentive, patient voice; non-potentive, patient voice; non-potentive, unaccusative; or adjectival.[1] With the exception of adjectivals, each of these functions of Subanon *mo-* corresponds with one of a cognate *ma-* in Tagalog. Himmelmann (2006) resolved the issue of describing this multifunctional prefix in Tagalog by showing that a single paradigm involving the Tagalog prefix *ma-* was insufficient to account for its multiple functions. He instead identified its role in two separate paradigms: statives and the potentive formation of dynamic predicates (itself in a paradigmatic relationship with non-potentive dynamics). With these separate paradigms, Himmelmann was able to systematically explain the behavior of *ma-*, which had seemed unpredictable to previous analysts.

Given the similarities between Subanon and Tagalog, a reasonable hypothesis would be that multiple paradigms, equivalent to the Tagalog potentive and stative, must be defined to describe the Subanon *mo-*. However, I show that this approach is not adequate for Subanon, as it would define more paradigms than necessary and miss an important generalization: that all predicates bearing affixes from Table 1 are associated with lower agentivity (Dowty 1991) or lower transitivity (Hopper & Thompson 1980) than the same predicates with volitional affixation.

Dowty (1991) defines thematic roles as cluster concepts, called Proto-Agent (henceforth, P-Agent) and Proto-Theme. The P-Agent role is associated with the following entailments: volition, sentience, causation, movement, and independent existence (Dowty 1991: 572-573). In this framework, an argument may lack some of the entailments associated with the P-Agent role, and cross-linguistically grammars can differ in terms of how many entailments must be met for an argument to be treated, grammatically, as an agent. Thus, in these terms, the subjects of *mo-*affixed predicates in Subanon uniformly have fewer of the entailments associated with the P-Agent thematic role, irrespective of other denotations or functions associatd with this prefix.

I argue that the most elegant analysis of Subanon takes this into account, such that the interpretation of the Subanon *mo-* is a function of the prefix's role in the non-volitional

[1] Although certainly intersecting with the paradigms described here in interesting ways, Subanon adjectivals are unfortunately outside the scope of the current paper.

paradigm *in interaction with* a predicate's argument structure. To that end, the relevant background information is briefly summarized in section 2. In section 3, data justifying the non-volitional paradigm in Table 1 is provided. Section 4 motivates an analysis of the Subanon data that favors unified paradigms, mediated by argument structure, and section 5 offers conclusions.

2 Background

2.1 Tagalog paradigms

The prefix *ma-* in Tagalog (and indeed in many Austronesian languages) has long been noted to be multifunctional, used with predicates denoting involutary actions, accidents, abilities, statives, and more. Himmelmann (2006) observed that previous descriptions were unable to reconcile these meanings in Tagalog as relating in a systematic way to a single prefix. The key insight that allowed Himmelmann to describe the data more effectively was that this prefix actually belongs to two separate paradigms. The first is the potentive formation of dynamic predicates. In this paradigm, the prefix *ma-* is used for patient voice predicates denoting an ability or an involuntary action of the genitive-marked argument (i.e., the argument marked by *ng*). Although the combination of involuntary action into the same category as ability may seem semantically incongruous to a speaker of a non-Austronesian language, these readings are often morphologically united in modern Philippine and Oceanic languages, such that only context resolves the ambiguity between the involuntary, accidental, and ability readings for such predicates. Therefore, it is desirable to treat this as a single function of the morpheme: i.e., potentive.

(1) na-dalá ko ang libró
POT.PV.REAL-carry 1SG.GEN FOC book
'I took the book by accident.' (Himmelmann 2006: 493)[2]

[2]Abbreviations used in glossing are as follows: AV actor voice, CV conveyance voice, DET determiner, EXC exclusive, FOC grammatical marker identfying focused argument, GEN grammatical marker identifying genitive argument, GV goal voice, INC inclusive, IRR irrealis, LV locative voice, OBL grammatical marker identifying oblique argument, NVOL non-volitional, PL plural, POT potentive, PV patient voice, REAL realis, SG singular, STAT stative, VOL volitional.

Example (1) shows a prefix from the potentive *ma-* paradigm, here the realis *na-*, translated with an accidental-event reading. The agent of this action *ko* '1SG' has less control over his or her action than an agent of the same verbal base with non-potentive affixation; i.e., it lacks the P-Agent entailment of volition described by Dowty (1991).

Additionally, the *ma-* prefix in Tagalog belongs to a second paradigm: statives. Himmelmann defined statives as "forms referring to states of affairs which principally exclude the involvement of an agent" (2006: 510). In syntactic terms, the argument to which the state of affairs is predicated is marked with the focus marker *ang* (nominative), and genitive arguments (marked with *ng*) are ungrammatical; an oblique (marked by *sa*) may be permitted. Tagalog statives contrast with dynamics, which allow the involvement of an agent, although that agent may lack the entailment of volition (i.e., potentives). In the patient voice of dynamics, the agent is expressed as the genitive argument. Previous analysts most likely missed this crucial distinction between potentive dynamics and statives, because in both cases *ma-* affixes to the same types of bases and so initially appears to have the same syntactic distribution. However, the difference in argument structure between dynamics and statives shows that their syntactic distribution is in fact distinct.

Himmelmann further showed that the semantic entailments of the *ma-* prefix differ between the dynamic and stative paradigms. In the dynamic verb paradigm, *ma-* is only used for potentive formations, denoting an accident or ability, and is therefore more semantically marked in comparison with non-potentive dynamics, which denote a wider array of events. For statives, the *ma-* paradigm is used as the default predicate conjugation and cannot be interpreted with a potentive meaning. Further, for statives, the forms prefixed with the irrealis *ma-* or realis *na-* (patient voice among dynamics) can be considered the default for statives, because these forms are possible for all statives, and voice alternations are rarely permitted. This contrasts with dynamic verbs, where no particular voice alternation can be said to be the default or unmarked form. Himmelmann (2006) presents this fact as further evidence in favor of distinguishing separate paradigms. Example (2) shows a stative with the realis *na-* from the stative paradigm.

(2) na-pipi siyá
STAT.REAL-dumb 3SG
'He got dumb.' (Himmelmann 2006: 491, citing Bloomfield 1917: 285/29)

The paradigms in which the *ma-* prefix appears in Tagalog are reproduced from Himmelmann (2006) in Table 2 (irrealis mood, perfective aspect only), including the dynamic non-potentive paradigm for comparison with the potentive. While the patient and actor voice potentive forms coincide with the equivalent voice of the statives, both locative and conveyance voice forms differ between potentives and statives; that is, these forms lack absolute morphological covariation, a crucial criterion for uniting morphological paradigms. This further supports the partitioning of the *ma-* prefix into separate paradigms.

Table 2: Paradigms involving the *ma-* prefix in Tagalog, irrealis mood, perfective aspect only, reproduced from (Himmelmann 2006: 517)

	Dynamic		**Stative**	
	Non-potentive	*Potentive*		
PV	-in	ma-	ma-	STAT
AV	-um-, mag-	maka-	(maka-)	STAT.AV
LV	-an	ma-...-an	ka-...-an	STAT.LV
CV	i-	ma-i-	ika-	STAT.CV

Table 2 further illustrates another compelling piece of evidence for splitting paradigms: the dynamics and statives are paradigmatically related to one another. That is, statives in many cases can employ dynamic, non-potentive affixation and take on an agent argument: the (patient voice) stative *ma-takot* 'be afraid' can take the form *t-um-akot* 'frighten' as in an actor voice dynamic pattern. The reverse is also true, in which typically dynamic predicates can employ stative morphology, rendering an agent argument ungrammatical.

Finally, Himmelmann also highlights the differing thematic roles for the arguments of dynamics and statives formed with *ma-* as evidence for separate paradigms. When a dynamic predicate is marked with the *ma-* prefix, the subject (focused argument) is a patient or theme, and the genitive is an agent or experiencer. For statives, on the other hand, the subject is a theme or experiencer, and a genitive argument is not permitted.

Table 3 provides a summary of the evidence Himmelmann (2006) gives that the *ma-* prefix participates in the two paradigms given in Table 2. This evidence is, essentially, of three different types: morphological (1-3), syntactic (4-5), and semantic (5-6). Importantly, each of these factors that Himmelmann considers points toward the same conclusion: the need to define two separate paradigms featuring *ma*-. The question one must always ask when multiple factors are used to define categories is, what if the evidence points in different directions? Which criteria should be decisive in defining morphological paradigms – the morphological, syntactic, or semantic?

Table 3: Evidence for split Tagalog potentive and stative paradigms (Himmelmann 2006)

	Evidence	**Conclusion supported**
1	The paradigms lack absolute morphological covariation.	split paradigms
2	The dynamic and stative paradigms are paradigmatically related to one another.	split paradigms
3	There is no default voice for dynamic predicates. For statives, the default is the *ma*-marked voice.	split paradigms
4	*Ma*-marked statives are ungrammatical with a genitive argument, while *ma*-marked potentives permit a genitive.	split paradigms
5	The subjects and genitive arguments of *ma*-marked dynamics have different thematic roles than the corresponding arguments of *ma*-marked statives.	split paradigms
6	*Ma*-marked dynamics are potentive formations, while *ma*-marked statives are not read as potentive.	split paradigms

Subanon *mo*- presents such a case. Based on the syntactic and semantic features of this prefix, multiple paradigms could be defined, as in Tagalog. However, based on morphological evidence, only one paradigm with *mo*- is necessary: non-volitional. I argue that, in defining morphological paradigms, only morphological evidence should be considered, and that syntactic-semantic incongruities should be resolved with a syntactic-semantic approach – in the case of Subanon, by explaining these patterns in terms of argument structure.

2.2 The volitional paradigm

Before turning to a discussion of the Subanon *mo*- and its paradigmatically related forms, it is necessary to first establish one paradigm in which this prefix does not participate:

the volitional paradigm (equivalent to the Tagalog dynamics), as shown in Table 4 and illustrated in examples (3)-(5) with the bases *bunu'* 'kill' and *bogoy* 'give.'

Table 4: The Subanon volitional paradigm

	IRR	REAL
PV	-on	-in-
AV	-um-	-um-in-
GV	-an	-in-. . . -an

(3) bunu'-on=ku og gotow koyon
kill-VOL.PV=1SG.GEN FOC person DET
'I will kill that person.'[3] (FM1-008-a.wav 00:01:37-00:01:42)[4]

(4) <m>unu' og polopanad nog manuk
<VOL.AV>kill FOC teacher GEN chicken
'The teacher will kill the chicken.' (FM1-026-b.wav 00:43:17-00:43:21)

(5) b<in>ogoy-an=ku og lokole=u nog lot
<VOL.REAL>give-GF=1SG.GEN FOC friend=my GEN knife
'I gave my friend a knife.' (FM1-013-a.wav 01:06:21-01:06:23)

The thematic roles of the arguments in these examples follow the general pattern of Philippine voice systems. In all cases, there is minimally an agent, who deliberately performs the action (e.g., 'the teacher' in (4)), and either a patient, which undergoes the action or a change of state ('that person' in (3)) or theme, which undergoes the action but not a change of state ('a knife' in (5)). The arrangement in which the patient or theme is marked with the focus marker *og*, as in (3), illustrates the patient voice; the arrangement in which the

[3]Because the nuances conveyed by Philippine voice systems are difficult to render in English, translations will be given in a neutral English word order, with the focus-marked element underlined in translation.

[4]All Subanon examples were provided by Sharon Estioca, a linguist and native speaker of the Malayal dialect of Subanon, in the context of a graduate-level field methods class at UHM between August 2015 and May 2016. Materials are archived at the *Kaipuleohone Language Archive*: `https://scholarspace.manoa.hawaii.edu/handle/10125/40825`. When a recording is available for an example sentence, a file name and start-finish timestamp (HH:MM:SS) are provided. When an example sentence is from a recorded text, the name of the text is given in this citation as well; all other examples are elicited. Subanon sentences without these indicators are from the author's notes of elicitation sessions that were not archived.

agent is the focused argument, as in (4), exemplifies the actor voice; and the arrangement in which the recipient is the focused argument, as in (5), is called the goal voice.

In these examples, the agent arguments are entailed as having volition: they act deliberately and are in control of the action. This paradigm is therefore labeled *volitional*, in contrast with the *non-volitional* paradigm that will be established in subsequent sections.

3 The Subanon non-volitional paradigm

Not all predicates in Subanon are affixed according to the paradigm shown in Table 4. A common alternate affix is *mo*-. Like its cognate in Tagalog, the Subanon *mo*- is associated with a variety of interpretations. It can be found attached to predicates that denote accidental events, abilities, acts of perception, emotional and physical states, locomotion, and properties. For each associated meaning, the realis prefix *mo*- alternates with the irrealis form *mi*-. In addition, for all meanings, *mo*- has the same allomorphs: a metathesized allomorph *om*- before labials (*mo+bogat* → *ombogat* 'heavy'), and the simplified *m*- with vowel-initial bases (*mo+init* → *minit* 'warm').

In this section, the two main interpretations of *mo*-marked predicates, potentive and unaccusative, are illustrated. The use of *mo*- with properties (i.e., adjectivals) and locomotion predicates is outside the scope of the present paper.

3.1 Potentives

In examples (6) and (7), the realis variant *mi*- is shown to receive a potentive reading. In (6), a genitive-marked agent accidentally commits the activity that affects the focused patient. That is, this form is in patient voice, but unlike the patient voice predicates of the volitional paradigm, as in examples (3) above, the agent argument lacks the entailment of volition. Similarly, in (7), the genitive actor is not a prototypical agent, but rather one who is ascribed an ability or who manages to commit the action in the face of some impediment.

These examples are glossed as "non-volitional," although at this point the given evidence only suggests that these are equivalent to the potentives in the Tagalog dynamic paradigms. The *mo*-prefix associated with this potentive reading alternates with other morphemes, *moko-* and the discontinuous *ko-/-an*, which rearrange the focus of arguments in parallel with the voice alternations established in the volitional paradigm (Table 4).

(6) mi-lisag=u nog gina'=u nog ponglisag=non
NVOL.PV.REAL-strike=1SG.FOC GEN mother=my GEN swatter=his/her
'My mother (accidentally) struck me with her swatter.'
(FM1-011-a.wav, 00:49:40-00:49:55)

(7) mi-tagu'=non na og lam nog gumpan sog bubu
NVOL.PV.REAL-put=3SG.GEN already FOC all GEN bait OBL fish.trap
'S/he was able to (or accidentally) put all the bait in the fish traps.'
(FM1-012-a.wav 02:08:49-02:08:52)

(8) moko-bunu'=u nog gotow
NVOL.AV.IRR-kill-1SG.FOC GEN person
'I will accidentally kill that person.'

(9) ko-bogoy-an=ku og lokole=u nog lot
NVOL.GV.IRR-give-GV=1SG.GEN FOC friend=my GEN knife
'I will accidentally give my friend a knife.'

Examples (8) and (9) are the equvalent of (4) and (5), but the action denoted by the predicate differs in that it was not committed intentionally. In (8), where the affix is the irrealis *moko-*, the agent is the focused subject, and the genitive argument is the patient. This example, then, shows the actor voice with a potentive interpretation. In (9), where the affix is the irrealis *ko-/-an*, the focused argument is the recipient, and there are two non-focused arguments, an agent and a theme; i.e., goal voice with a potentive interpretation.

Table 5: The Subanon non-volitional paradigm

	IRR	REAL
PV	mo-	mi-
AV	moko-	miko-
GV	ko-...-an	ki-...-an

A paradigm involving *mo-* can now be established for three voice types and two moods: the paradigm given in Table 1, reproduced here in Table 5. This paradigm illustrates that a form prefixed with *mo*- relates to one prefixed with *mi*- in the same way that a *moko*-prefixed form relates to a *miko*-prefixed one, and so on. It does not, however, guarantee that a form bearing all such prefixes exists for every base. While both realis and irrealis mood exist for all bases used as potentives in Subanon, not all voice alternations are possible. As with other Philippine languages, the voice alternations available for a given lexical item in Subanon depend on the meaning of the base and are highly conventionalized.

Thus far, the morphology in Table 5 has only been shown to be associated with a potentive reading. However, the term "non-volitional" is used because, as will be shown later, this *mo*- paradigm more generally functions to cancel out the volitionality entailment of the most agent-like argument (i.e., the one bearing the most P-Agent entailments) .

The parallelism between the potentive examples (8) and (9) and the volitional examples (4) and (5) further suggests that these paradigms bear a paradigmatic relationship to one another. That is, *munu'* 'kill:VOL.AV.IRR' in example (4) differs from *mokobunu'* 'kill:NVOL.AV.IRR' in only with respect to the agent's control over the action; i.e., the former receives a non-potentive interpretation and the latter potentive.

Table 6: Combined volitional and non-volitional paradigms

	Volitional		**Non-volitional**	
	IRR	REAL	IRR	REAL
PV	-on	-in-	mo-	mi-
AV	-um-	-um-in-	moko-	miko-
GV	-an	-in-...-an	ko-. . . -an	ki-. . . -an

In fact, all bases that can take some affix from the volitional paradigm in Table 4 can also take the equivalent affix from the paradigm in Table 5, where the only difference will be whether the agent is entailed as acting with volition. Thus, these two paradigms can be combined into a greater paradigm as in Table 6.[5] This arrangement closely resembles the non-potentive/potentive paradigms of dynamics given by Himmelmann (2006) for Tagalog. The next question is whether Subanon unaccusatives, which are similar to Tagalog

[5]Subanon has another set of affixes (*mog*-/*mig*-, *pog*-/*pig*-) used for volitional voice alternations. However, it is currently unclear how these affixes differ from those given in the chart; i.e., if these are conjugation classes or derivational morphemes.

statives, neatly fit in the paradigm in Table 6, or whether an additional paradigm needs to be defined, like the stative paradigm Himmelmann defined for Tagalog.

3.2 Unaccusatives

In examples (10) and (11), the prefix *mi-* (i.e., the realis counterpart of *mo-*) is shown with a predicate denoting a physical or emotional state: sleep in (10) and fear in (11). In (10), the focused argument is a theme, and a genitive argument (*nog*) would be ungrammatical. In (11), the focused argument is an experiencer, and the genitive argument is the stimulus triggering the fear. I call formations of this type "unaccusatives."

(10) gobi koyon ditu' ami mi-tulug sog gurnu
night DET there 1PL.EXCL.FOC NVOL.REAL-SLEEP OBL copra.oven
'That night, we slept in the room where copra is processed.'
(Unnamed text about bad weather, FM1-024.pdf)

(11) mi-ondok=u nog gotow koyon
NVOL.REAL-afraid=1SG.FOC GEN person DET
'I am afraid of that person.'

The presence of a genitive argument in (11) makes it appear to be a two-place predicate, in which case the term "unaccusative" would be inappropriate. However, example (12) confirms that the arguments in this construction represent a lower degree of transitivity that the two-place predicates shown in previous sections. This example shows the same verb root *(g)ondok* 'afraid', but the arrangement of arguments is reversed: the stimulus is now the focused element, and the experiencer is genitive. Importantly, the morphology on the verb is the same affixation established in section 3.1 for *goal* voice predicates, suggesting that the stimulus argument in (11) was in fact an oblique. Thus the predicate *(g)ondok* 'afraid' is not transitive and can be grouped with the other unaccusatives.

(12) og kukuk k-ondok-an=nami
FOC kukuk NVOL.GV.IRR-afraid-GV=1PL.EXCL.GEN
'We are afraid of the *kukuk*.'[6] (Text: *Og kukuk*. FM1-029.wav 00:03:32-00:03:34)

[6] In Subanon lore, a *kukuk* is a creature living in the mountains of Zamboanga. It has the ability to transform its face to look like that of a child's parent, in order to lure that child to its lair.

Some forms that express an emotional or physical state with the patient voice *mo-/mi-* affixation can appear with the actor voice *moko-/miko-* prefix of the non-volitional paradigm. However, the difference between these forms is not a shift in voice, as with other bases alternating between the same affixes. Rather, the difference between a *mo*-prefixed state as in (13) and a *moko*-prefixed state as in (14) is the interpretation of the act as non-potentive or potentive. The argument structure of *motulug* 'will sleep' and *mokotulug* 'able to sleep' remains the same: only one argument is permitted, the focused theme, and a genitive argument is ungrammatical in both cases. For some bases of this type, a *moko-/miko-* form would be ungrammatical, likely for semantic reasons (e.g., **mokondok* 'able to be afraid,' **mokogutom* 'able to be hungry').

(13) mo-tulug=u bombus
NVOL.PV.IRR-sleep=1SG.FOC later
'I will sleep later.'

(14) moko-tulug=u bombus
NVOL.AV.IRR-sleep=1SG.FOC later
'I will be able to sleep later.'

Thus, Subanon unaccusatives are potentially compatible with all affixes in the non-volitional paradigm. However, the interpretation of unaccusatives with these affixes differs, presenting a problem for paradigmatic analysis. Where the relationship between *mo-* vs. *moko*-affixed predicates in section 3.1 was a shift in voice, the same relationship among unaccusatives was instead a shift in volitionality, where only the *moko-* form receives a potentive reading. This disparity is incompatible with the paradigms in Table 6.

As explained in section 2.1, precisely this sort of incongruity in Tagalog led Himmelmann (2006) to propose a separate stative paradigm. Therefore, it is worth considering what such an arrangement would look like in Subanon. This split-paradigm approach, with differentiated non-volitional (i.e., potentive) and unaccusative (stative) paradigms, is shown in Table 7, mirroring Himmelmann's treatment of Tagalog (cf. Table 2).

By partitioning the paradigms in this way, it is now possible to account for Subanon unaccusatives: they simply belong to a different paradigm. Indeed there is some potentially convincing evidence in favor of this approach. As in Tagalog, these paradigms in Subanon

Table 7: An alternate arrangment of Subanon paradigms

	Volitional	Non-volitional	Unaccusative	
PV	-on	mo-	mo-	PV
AV	-um-	moko-	moko-	PV.POT
GV	-an	ko-...-an	ko-...-an	GV

appear to be in paradigmatic relationship with one another, as demonstrated in examples (15) - (17). Example (15) shows a predicate *lonod* 'sink,' bearing an affix from the non-volitional paradigm, and a genitive argument would be ungrammatical. This predicate is therefore not a potentive, since patient voice potentives allow a genitive agent (deprived of volition). Rather, it must be an unaccusative. In comparison, the same predicate is shown in (16) with the same non-volitional affixation, receiving a potentive reading. In (17), the predicate is minimally different in that it now bears volitional affixation, and accordingly, it receives a non-potentive reading.

(15) mi-lonod og batu
NVOL.REAL-sink FOC stone
'The stone sank.'

(16) mi-lonod nog polonoda' kitu' og bolangoy=non
NVOL.REAL-sink GEN fisherman DET FOC boat=his/her
'The fisherman accidentally sank his/her boat.'

(17) l<in>onod nog polonoda' kitu' og bolangoy=non
<VOL.PV.REAL>sink GEN fisherman DET FOC boat=his/her
'The fisherman sank his/her boat (on purpose).'

However, even though this approach would account for the data, I find the paradigms in Table 7 uncompelling for a critical reason: the affixation of non-volitionals is identical to that of unaccusatives. The split-paradigm approach, while amply justified in Tagalog, fails to account for the absolute morphological covariation among non-volitionals and unaccusatives in Subanon. A different approach is needed that will both capture these generalizations and still account for the data. Such a solution is proposed in section 4.

4 Interpreting non-volitionals using argument structure

In Table 3 in section 2.1, six criteria were outlined to support the partitioning of Tagalog *ma-* into two separate paradigms, the potentives and the statives. This evidence was of three different types: morphological (1-3), syntactic (4-5), and semantic (5-6). In the case of Tagalog, crucially, the evidence, regardless of its domain, always pointed toward the same conclusion: to support splitting the paradigms. In Subanon, the same evidence for unifying vs. splitting paradigms was presented in section 3 and is summarized in Table 8.

Table 8: Evidence for splitting vs. unifying Subanon paradigms with *mo-*

	Evidence	**Conclusion supported**
1	There is absolute morphological covariation in all forms.	**united paradigm**
2	There is a paradigmatic relationship among volitionals, non-volitionals, and unaccusatives.	split paradigms
3	There is no default voice for volitional and non-volitional predicates. For unaccusatives, the default is the *mo*-marked voice.	split paradigms
4	*Mo*-marked unaccusatives are generally ungrammatical with a genitive argument. *Mo*-marked non-volitionals permit a genitive.	split paradigms
5	The subjects and genitive arguments of *mo*-marked volitionals have different thematic roles than the corresponding arguments of *mo*-marked unaccusatives.[7]	split paradigms
6	Non-volitionals marked with *mo-* receive potentive interpretations, while *mo*-marked unaccusatives are not read as potentive.	split paradigms

Unlike in Tagalog, the Subanon evidence did **not** always support the same conclusion. The majority of the evidence pointed toward splitting the paradigms, with a separate non-volitional and unaccusative as in Table 7. Only the morphological criterion favors unifying the paradigms, but it is a compelling piece of evidence: there is absolute morphological covariation among the affixation used for non-volitional and unaccusative predicates.

How could a child speaker even acquire two separate paradigms without some kind of morphological differentiation? In Tagalog, these two paradigm sets were differentiated by the fact that conveyance voice and locative voice forms differed morphologically, making the paradigms learnable. In Subanon, however, conveyance and locative voice are collapsed into goal voice, and the forms are not morphologically distinct. An acceptable model of Subanon must capture this morphological generalization.

Moreover, the split paradigm approach misses yet another generalization, not stipulated in Table 8: the semantic generalization that all predicates with *mo-* and its paradigmatically related affixes involve an argument structure with fewer P-Agent entailments (to use terminology from Dowty (1991); cf. recent work by Foley (2005)). This can be illustrated with the examples (15)-(17) in section 3.2. In these examples, the same root *lonod* 'sink' showed a gradient with respect to agentivity, conceptualized in terms of how many P-Agent entailments the arguments have. The form *linolod* 'sank' with a volitional infix in (17) exhibited the highest degree of agentivity, where an agent (the fisherman) deliberately sank his or her boat; that is, the fisherman was clearly compatible with the volitional entailment. On the low end, no arguments of the unaccusative form *milonod* 'sank' in (15) had any P-Agent entailments, the stone's sinking being an unremarkable consequence of physics. The identical form in (16), *milonod*, fell in the middle of the cline and receives a potentive interpretation; an agent is stated, but lacks the volition entailment. In essence, along a scale of agentivity, the Subanon grammar defines two points toward the lower end, the first receiving a potentive reading and the lowest being an unaccusative, and it so happens in Subanon that both of these levels are marked with identical morphology.

Based on these generalizations, the united non-volitional paradigm given in Table 6 better describes the Subanon data. The ability to predict the meaning of an affixed form is retained for Subanon by turning to argument structure to mediate between the potentive vs. unaccusative interpretations of non-volitionals. To make that clear, I offer operational definitions of the the terms "volitional," "non-volitional," and "actor voice" in terms of the P-Agent entailments associated with the most agent-like argument in 1. Drawing on these definitions, the two possible interpretations of a predicate affixed from the non-volitional paradigm, potentive or unaccusative, are predictable from argument structure. In fact, these interpretations can now be defined quite simply, as in point 2.

1. Morphological definitions

    (a) **non-volitional** = no arguments of the predicate are entailed to act with volition.

    (b) **volitional** = the argument of the predicate that bears the most other P-Agent entailments is entailed also to act with volition.

    (c) **actor voice (AV)** = the argument with the most P-Agent entailments is the grammatical subject marked by *og* (FOC) or a focused pronoun.

2. Definitions of interpretations

    (a) **unaccusative** = a state of affairs in which no participants have enough P-Agent entailments to be treated as an agent (i.e., the argument structure excludes an agent).

    (b) **potentive** = a state of affairs in which the agentive participant lacks the entailment of volition.

The definition of unaccusatives given in 2a closely resembles the Himmelmann's definition of statives for Tagalog. It also becomes clear by comparing these definitions why an unaccusative naturally pairs with non-volitional morphology. A potentive predicate, on the other hand, has the argument structure of volitionals (i.e., it includes an agent) but the morphology of non-volitionals, putting this interpretation somewhere in the middle.

Using these definitions, it is now possible to account for the idiosyncracies that countered the united paradigm approach. Examples (18) and (19) (reproduced from (7) above) show the same base, *tagu'* 'put' with volitional actor voice affixation and non-volitional patient voice affixation. By the definitions above, then, *tumagu'* 'will put' should be understood to mean that argument of the predicate that bears the most other P-Agent entailments (the first person actor) is entailed also to act with volition, and that the argument with the most P-Agent entailments is the grammatical subject marked by the focused pronoun *u* '1SG.' This in fact bears out in the given example. In contast, *mitagu'* 'was accidentally put' in (19) should be understood to mean that no arguments of the predicate are entailed to act with volition, which again is borne out in the potentive interpretation of this example.

(18) t<um>agu'=u sala pilak sog kansia=u
<VOL.AV>put=1SG.FOC one peso OBL piggy.bank=my
'I will (intentionally) put one peso in my piggy bank.'

(19) mi-tagu'=non na og lam nog gumpan sog bubu
NVOL.PV.REAL-put=3SG.GEN already FOC all GEN bait OBL fish.trap
'He/she was able to to (or accidentally) put all the bait in the fish traps.' (reproduced from (7))

The crucial test is whether these definitions account for the idiosyncratic behavior of unaccusatives. The interpretations of the forms with the root *tulug* 'sleep' in examples (20) and (21) (reproduced from (13) and (14) above) vacuously satisfy the definitions above. That is, in example (20), it is true that the most agent-like argument lacks volition, in that there is only one argument, which bears very few P-Agent entailments, excluding volition. In (21), it is similarly true that the argument with the most P-Agent entailments is the subject marked with *og*. Both of these examples can be captured by the defintion for unaccusatives as a state of affairs in which no participants have enough P-Agent entailments to be treated as an agent (i.e., the argument structure excludes an agent).

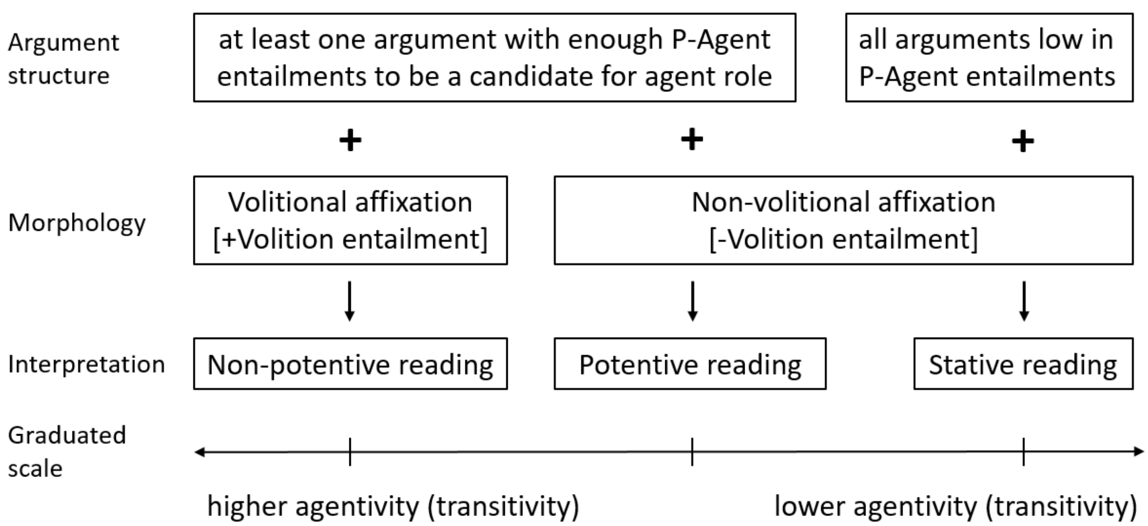
(20) mo-tulug=u bombus
NVOL.PV.IRR-sleep=1SG.FOC later
'I will sleep later.'

(21) moko-tulug=u bombus
NVOL.AV.IRR-sleep=1SG.FOC later
'I will be able to sleep later.'

Therefore, these examples are not inconsistent with the definitions I have proposed. It remains unclear whether these definitions offer a satisfying explanation for why the *mo*-prefixed form in (20) does not receive a potentive reading, yet the actor voice form, prefixed with *moko-*, in (21) does. With the current limited availably of Subanon data, it may not be possible at this time to offer a truly satisfying explanation of this phenomenon.

Nonetheless, although this account of *moko-* is not ideal, the united paradigm approach, combined with these definitions referencing argument structure, appears to be operationally suitable for Subanon and in fact has a lot of predictive power. The multiple functions of *mo-* are fittingly united by a single, consistent function: the role of *mo-* and other affixes from the non-volitional paradigm is, quite simply, to cancel out the P-Agent entailment of volition. This is schematized in Figure 1. A scale of this sort could be imagined to be universal across languages that show some grammatical sensitivity to agency, while the stopping points along the scale would be language-specific. The grammar of Subanon has three stopping points along this cline—volitional, potentive, and stative—which it encodes morphologically with the aid of two sets of affixes.

Figure 1: Schematization of the interaction of argument structure and morphology (agentivity scale)



The function of Subanon non-volitional affixes such as *mo-*, in terms of this diagram, is to move a predicate rightward along the scale, decreasing agentivity. Likewise, the function of Subanon volitional affixes triggers a move to the left of the scale, increasing agentivity. The scale could similarly be interpreted in terms of higher or lower transitivity (cf. Hopper & Thompson (1980)). A move rightward would then be termed 'detransitivization,' while a leftward move along the scale would be 'transitivization.'

A desirable consequence of the schema in Figure 1 is that some lexical items can be associated with a 'default' position along this scale. Some predicates, like *bunu'* 'kill,' are more typically expressed with a volitional agent. As such, their default position is on the left end of the scale, and a form with non-volitional affixation (i.e., *ombunu'* 'accidentally kill') involves a step to the right on the scale and a lower degree of agentivity (or transitivity). An unaccusative version of this predicate (meaning, perhaps, 'to get killed') would then bear *mo*- affixation and have only one argument. However, it is not clear an unaccusative form indeed exists in Subanon. If it were necessary to exclude such a form, it could be assumed that *bunu'* 'kill' is stipulated in the lexicon as having only one available argument structure, which includes an agent. Thus, the unaccusative would always be ruled out for this predicate, because a form of *bunu'* with the argument structure necessary for a stative interpretation would not exist in the lexicon.

On the other end, predicates like *gutom* 'be hungry' or *(g)ondok* 'be afraid' semantically exclude an agent, and their default position can be assumed to be at the right end of the graduated scale; i.e., unaccusatives. In some cases, predicates of this type can take affixation from the volitional paradigm, which has the effect of introducing an agent. For example, the *mo*- in *mosasow* 'be annoyed' can be replaced with volition affixation to achieve the meaning 'to annoy.' Unfortunately, it is impossible to identify the "default" position of most predicates along the agentivity scale, until a large corpus of Subanon is collected, from which word frequencies could be gathered.[8]

5 Conclusions

It is now clear that a description of the *mo*- prefix by itself, without illustrating its position in terms of paradigms, would be insufficient. The behaviors of some predicates are unpredictable enough that, if we started from the assumption that all instances of *mo*- are instances of the same prefix performing the same function, we would be led astray, and the task of defining the function of the *mo*- prefix, without teasing apart its separate meanings and functions, would be nearly impossible.

However, unlike Tagalog, in Subanon there is not enough justification for a separate unaccusative (or stative) paradigm, with different functions and different morphemes, in contrast to the potentive mood paradigm. The correct interpretation of such predicates

[8]A valid question that arises from this discussion is whether any predicates assume a default position in the *middle* of the agentivity scale; i.e., a default potentive predicate. This appears to be the case for acts of perception, which are outside the scope of the present paper.

is mediated by argument structure; that is, speakers identify stative interpretations by the omission of any agent-like argument and potentives by the agent's inclusion.

This differs from Himmelmann's analysis of the Tagalog cognate *ma-*. In Tagalog, Himmelmann showed that the potentive paradigm is separate from, but paradigmatically related to, a paradigm of statives. The Tagalog stative paradigm nonetheless includes the *ma-* prefix, but behaves differently, both morphologically and syntactically. In considering multiple types of evidence, each factor pointed in the same direction: to split the paradigms. Thus, defining separate potentive and stative paradigms in Tagalog is likely the optimal analysis.

The same cannot be said of Subanon, where unaccusatives bear the same affixation as other potentives in all voice arrangements. Thus, while the same syntactic and semantic criteria as Himmelmann used in Tagalog pointed toward splitting paradigms in Subanon, the morphological evidence, crucially, pointed in the opposition direction: toward unifying the potentive and stative interpretations into a single non-volitional paradigm.

This analysis highlights the risk inherent in using multiple sources of evidence in defining categories. When all the evidence converges, having multiple sources strengthens the analysis, as in Himmelmann's treatment of Tagalog. However, when the different factors considered suggest conflicting conclusions, one must ask which type of evidence should be considered decisive. The case of Subanon has shown that, if the categories being defined are morphological paradigms, then morphological evidence must be the deciding factor. Conflicting evidence from the domains of syntax and semantics is best resolved using syntactic or semantic strategies: in this case, mediating between the potentive and stative interpretations of *mo-* with argument structure. If syntactic and semantic factors had been used to justify these morphological categories, in the case of Subanon, we would have over-differentiated the paradigms and missed a generalization or two.

References

DOWTY, DAVID. 1991. Thematic proto-roles and argument selection. *Language* 67(3), 547–619. doi:10.2307/415037.

FOLEY, WILLIAM. 2005. Semantic parameters and the unaccusative split in the Austronesian language family. *Studies in Language* 29(2), 385–430. doi:10.1075/sl.29.2.05fol.

HIMMELMANN, NIKOLAUS P. 2006. How to miss a paradigm or two: Multifunctional *ma*- in Tagalog. *Catching language: The standing challenge of grammar writing*, ed. by Felix K. Ameka, Alan Charles Dench, and Nicholas Evans, 487–526. Berlin ; New York: M. de Gruyter.

HOPPER, PAUL J., and SANDRA A. THOMPSON. 1980. Transitivity in grammar and discourse. *Language* 56(2), 251–299. doi:10.2307/413757.

Two types of preverbal object movement and duration/frequency phrases in Mandarin Chinese

HSIN-LUN HUANG
*University of Massachusetts Amherst**

1 Introduction

In Mandarin Chinese (henceforth Mandarin), there are (at least) two types of object displacement to the preverbal position. The first type involves preposing a BARE object with no additional morpheme accompanying the preposed object (Ernst and Wang 1995, Paul 2002, 2005, Badan 2008, among others). The preverbal position the object moves into is defined as the position preceding the main verb (with apectual marking) and following the subject[1]. For convenience, I will term object-preposing of this type, as in 1, as TYPE I.

(1) Lisi **zhe-ben shu** nian-le san tian/ci – Type I
Lisi this-CL book read-ASP three day/time
'Lisi read this book for three days/three times.'

The second type, which I will term as TYPE II as in 2, involves a preverbal object PLUS a verb copy that precedes the object[2]. Type II is generally known as the verb-doubling construction in Mandarin and has received extensive analyses (Huang 1982, Cheng 2007, Tieu 2008, among others).

(2) Lisi nian **zhe-ben shu** nian-le san tian/ci – Type II
Lisi read this-CL book read-ASP three day/time
'Lisi read this book for three days/three times.'

*I would like to thank Kyle Johnson, Rajesh Bhatt, and Seth Cable for their valuable comments and suggestions throughout this project. I would also like to thank the audience at UMass Syntax Workshop, Spring 2018, for their valuable feedback and suggestions, and Rong Yin for her judgements on the Mandarin examples that helped shape the analysis. Of course, any errors that remain would be mine.

[1]Mandarin is an SVO language, so the object preceding the main verb is a more marked word order that generally requires some prompt for its occurrence.

[2]In the regard of Type II, calling the object preverbal is a little misleading, since it precedes, as well as follows the verb. For reasons to be clear later, the 'preverbal' position is defined with respect to the verb with aspectual marking. Therefore, according to that definition, the object in Type II IS preverbal.

Both Type I and II can be found in the presence of a duration/frequency adverbial phrase (abbreviated as DFP)[3]. Duration/frequency adverbials, such as *san tian/ci* 'for three days/three times' as in 1 and 2, occur post-verbally in Mandarin.

(3) Lisi (*san tian/ci) ku-le (san tian/ci)
Lisi three day/time cry-ASP three day/time
'Lisi cried for three days/three times.'

Traditionally, Type I and II are thought to be independent constructions that are subject to their own properties and restrictions. However, I will argue that Type II in the presence of DFPs cannot assume the previous analyses proposed for Mandarin verb-doubling. And by comparing the preverbal object in Type I to sentence-initial objects in Mandarin, I will show that the former is in the topic position in the middle field of Mandarin syntax, following the functional hierarchy proposed in Paul 2002, 2005 and Badan 2008. By further showing that Type II is subject to the same functional hierarchy, I will argue for a unified analysis for Type I and II that operates under one general rule of thumb: Sentence-internal topicalization.

The remainder of this paper is organized as follows. Section 2 looks at the distributions of Type I and II and shows that there is tremendous distributional overlap between the two movement types. Section 3 briefly reviews the literature on the patterns of Type I and II and points out the relevant issues with regards to the previous analyses. A unified analysis for the two movement types will be proposed in section 4, followed by the supporting predictions of word orders involving Mandarin ditransitive verbs in section 5. Finally, section 6 concludes the paper.

2 Distributions of Type I & II

As have been seen in 1 and 2, Type I and II are found in co-occurrence with DFPs. In addition, they are also found in the following constructions.

[3]As mentioned in fn. 1, Type I and II are more marked word orders in Mandarin that call for a prompt. And the presence of a DFP can be viewed as one of the prompts that enable Type I and II. As will be seen soon, there are cases other than that of DFPs where Type I and II are also found. More detailed discussions will be dedicated to those cases when we scrutinize Type I and II in section 2.

(4) a. Lisi (nian) **zhe-ben shu** nian-de [$_{\text{AdvP}}$ hen kuai]
Lisi read this-CL book read-DE very fast
'Lisi read this book very fast.'

b. Lisi (nian) **zhe-ben shu** nian-de [$_{\text{RsltP}}$ hen lei]
Lisi read this-CL book read-DE very tired
'Lisi got tired from reading this book.'

The construction in 4a involves the presence of a post-verbal manner adverbial phrase (abbreviated as AdvP), whereas that in 4b is a resultative construction in Mandarin that involves a post-verbal resultative phrase (abbreviated as RsltP) (Huang 1988, Cheng 2007, among others). Both of these constructions have in common the functional morpheme *de* attaching to the main (second) verb that is analyzed by many to be the introducer of the post-verbal phrases[4,5] (Huang 1988, Sybesma 1999, Cheng 2007, Tieu 2008, among

[4]Whether this *de* is the same in both constructions is up for debate, due to the obvious meaning differences between the two constructions. However, the *de*'s in the two constructions have the same written form, and the elements introduced by *de*, i.e. the AdvP and RsltP, occupy the same post-verbal position, as shown by their inability of co-occurrence.

(i) a. *Lisi (nian) zhe-ben shu nian-de [$_{\text{AdvP}}$ hen kuai] [$_{\text{RsltP}}$ hen lei]
Lisi read this-CL book read-DE very fast very tired
'Lisi read this book very fast and got tired as a result.'

b. *Lisi (nian) zhe-ben shu nian-de [$_{\text{RsltP}}$ hen lei] [$_{\text{AdvP}}$ hen kuai]
Lisi read this-CL book read-DE very tired very fast
'Lisi got tired from reading this book very fast.'

This observation suggests that the *de*-induced post-verbal adverbial and resultative phrases should be at least structurally analyzed on a par; thus, it is not surprising that Type I and Type II pattern the same in one if they pattern the same in the other. Of course, the introduced elements in the two constructions are modifying different things, i.e. the manner adverbial is modifying the verbal predicate, yet the resultative phrase is predicated of the subject. This will be seen as the morpheme *de*, being a functional element in the syntactic clausal spine, structurally selecting for two different types of complements, i.e. the AdvP and the RsltP, from within which the modificational differences arise. That is, the modificational differences are taken to result from the internal structures of the *de*-selected phrases, but not from the syntax of *de* per se.

[5]*De* and aspectual marking are mutually exclusive on the main verb (Gouguet 2006), (i), and the occurrence of DFPs is independent of *de*, (ii).

(i) a. Lisi (nian) zhe-ben shu nian-de-**(*le)** [$_{\text{AdvP/RsltP}}$ hen kuai/lei]
Lisi read this-CL book read-DE-ASP very fast/tired
'Lisi read this book very fast.'/'Lisi got tired from reading this book.'

b. Lisi (nian) zhe-ben shu nian-**(*le)**-de [$_{\text{AdvP/RsltP}}$ hen kuai/lei]
Lisi read this-CL book read-ASP-DE very fast/tired
'Lisi read this book very fast.'/'Lisi got tired from reading this book.'

others). And clearly, both Type I and II are viable in these constructions, as suggested by the optionality of a verb copy in front of the preverbal object.

We have defined Type I and Type II movement as preposing the object to the position immediately preceding the main verb, with or without an accompanying verb copy. In fact, in all of the constructions above, the DFP, adverbial, and resultative construction, the object can also get preposed to the sentence-initial position, with or without a verb copy.

(5) a. (Nian) **zhe-ben shu**, Lisi nian-le [$_{DFP}$ san tian/ci]
read this-CL book Lisi read-ASP three day/time
'Lisi read this book for three days/three times.'

b. (Nian) **zhe-ben shu**, Lisi nian-de [$_{AdvP/RsltP}$ hen kuai/lei]
read this-CL book Lisi read-DE very fast/tired
'Lisi read this book very fast.'/'Lisi got tired from reading this book.'

The sentence-initial position is the canonical topic position in Mandarin. The fact that the preposed object in 5 can be accompanied by a verb copy suggests that they form a constituent ([V + Obj.]) that has undergone sentence-initial topicalization. If we assume that the accompanying verb copy is part of the result of Type II movement, then it follows that Type II is moving not only the object, but a verbal constituent containing the object, to the preverbal position[6]. And we will argue that the preposed bare object under Type I occupies the same position by showing that the bare object and the verbal constituent pattern the same in their subjectivity to the sentence-internal functional hierarchy.

2.1 Type-II-specific properties

Before we move on to reviewing the previous analyses for Type I and II, there are some Type-II-specific properties worth mentioning that will help shape our final analysis. First, the first verb (V$_1$) has to be an identical copy to the main verb (Huang 1982, among others). Synonyms do not work.

(ii) Lisi (nian) zhe-ben shu nian-le/***-de** [$_{DFP}$ san tian/ci]
Lisi read this-CL book read-ASP/DE three day/time
'Lisi read this book for three days/three times.'

[6]It explains why we take the object to be in the preverbal position, despite its following a verb copy.

(6) a. *[$_S$ (**<u>Du</u>** zhe-ben shu), Lisi (**<u>du</u>** zhe-ben shu) **<u>nian</u>**<u>-le</u> san tian/ci]
read this-CL book Lisi read this-CL book read-ASP three day/time
'Lisi read this book for three days/three times.'

b. *[$_S$ (**<u>Du</u>** zhe-ben shu), Lisi (**<u>du</u>** zhe-ben shu) **<u>nian</u>**<u>-de</u> hen kuai/lei]
read this-CL book Lisi read this-CL book read-DE very fast/tired
'Lisi read this book very fast.'/'Lisi got tired from reading this book.'

Second, V_1 has to be BARE (Li and Thompson 1981, among others).

(7) a. (<u>Nian-***le**</u> zhe-ben shu), Lisi (<u>nian-***le**</u> zhe-ben shu) <u>nian-**le**</u> san tian/ci
read-ASP this-CL book Lisi read-ASP this-CL book read-ASP 3 day/time
'Lisi read this book for three days/three times.'

b. (<u>Nian-***de**</u> zhe-ben shu), Lisi (<u>nian-***de**</u> zhe-ben shu) <u>nian-**de**</u> hen kuai/lei
read-DE this-CL book Lisi read-DE this-CL book read-DE very fast/tried
'Lisi read this book very fast.'/'Lisi got tired from reading this book.'

In the examples above, Type II and its sentence-initial counterpart are combined in all of the constructions previously mentioned, i.e. with DFPs, manner adverbials, and resultatives, where they show the same behavior with respect to the properties.

These properties are revealing in terms of the underlying mechanism of Type II. The first property suggests a copy-and-movement analysis for Type II: The only way to ensure the identicalness between the two verbs is by movement of the VP, since if the preverbal (or sentence-initial) VP were base-generated, requiring the verb identicalness between the preposed and the main VP would be an ad hoc stipulation.

And the second property shows that the copying mechanism probably applies to verbal constituents of a certain size. That is, assuming Aspect being a head (Asp^0) in the syntactic clausal spine and the main verb undergoing cyclic head-movement to Asp^0, then copying is not applicable to AspP, whose head would be the aspectually marked main verb.

3 Previous analyses on Type I & II

As mentioned before, Type I and II are generally considered as independent constructions that are derived via different syntactic operations. Type I is widely analyzed as focus movement of the object to the sentence-internal focus position (Ernst and Wang 1995, Shyu 1995, among others). However, Paul (2002, 2005) argues that the preverbal object under Type I behaves more like topics than foci, by comparison to sentence-initial topics[7]. She proposes the following functional hierarchy for Mandarin, where there is a sentence-internal topic position strictly above the sentence-internal focus position[8].

(8) CP > TopicP* > 'even' FocusP > IP > **inner TopicP** > 'even' FocusP[9] > *v*P

The functional hierarchy above is based on the syntactic interaction between the preposed bare object and the 'even' construction in Mandarin, the *lian... dou/ye* construction. The 'even' construction consists of two morphemes. One is the morpheme *lian* that roughly translates to 'even' in English and focuses any constituent immediately following it, forming a focused phrase. The other is the morpheme *dou/ye* that roughly translates to 'all/also'. It is obligatorily present and follows the even-focused phrase[10]. Given a ditransitive sentence in 9a as a baseline, we can *even*-focus, for instance, the indirect object (IO). And the focused IO can occur either preverbally or sentence-initially, 9b.

(9) a. Lisi jiao-le [Mali]^{IO} san tian [wuli]^{DO} (IO/DO = (in)direct object)
Lisi teach-ASP Mary 3 day physics
'Lisi taught Mary physics for three days.'

b. (Lian [Mali]^{IO}), Lisi (lian [Mali]^{IO}) dou jiao-le san tian [wuli]^{DO}
even Mary Lisi even Mary all teach-ASP 3 day physics
'Lisi taught even Mary physics for three days.'

Using the 'even' construction, Paul (2002) shows that the preverbal object in Type I is subject to the hierarchy as a topic in having to be higher than the *even*-focused phrase.

(10) a. Lisi ✓**(wuli)** lian Mali ***(wuli)** dou jiao-le san tian
Lisi physics even Mary physics all teach-ASP three day

[7]For a detailed argument for the preverbal object's similarities to sentence-initial topics and dissimilarities to (sentence-initial and preverbal) foci, please refer to Paul 2002, 2005.

[8]The range of being sentence-internal is bounded by the IP, as underlined in 8.

[9]The Foc(us)P here specifically refers to those phrases circum-marked by *lian... dou/ye*. In other words, we are making a distinction that a preverbal focused item necessarily comes with *lian... dou/ye* and negating the possibility of a sentence-internal, syntactically bare FocP.

[10]According to Paul 2002, 2005, *dou* 'all' and *ye* 'also' are interchangeable in this construction. We will not be too concerned about the analysis of the *lian... dou/ye* construction. The point is simply to demonstrate the hierarchy in 8 with concrete examples.

'Lisi taught even Mary physics for three days.'

b. ✔**(Wuli)**, lian Mali, ***(wuli)**, Lisi dou jiao-le san tian
physics even Mary physics Lisi all teach-ASP three day
'Lisi taught even Mary physics for three days.'

I show that the [V + Obj.] constituent in Type II patterns the same as the preverbal bare object in its subjectivity to this hierarchy[11].

(11) a. Lisi ✔**(jiao wuli)** lian Mali ***(jiao wuli)** dou jiao-le san tian
Lisi teach physics even Mary teach physics all teach-ASP three day
'Lisi taught even Mary physics for three days.'

b. ✔**(Jiao wuli)**, lian Mali, ***(jiao wuli)**, Lisi dou jiao-le san tian
teach physics even Mary teach physics Lisi all teach-ASP three day
'Lisi taught even Mary physics for three days.'

This observation points to the preverbal items in Type I and II (the object and [V + Obj.] constituent, respectively) being very likely in the same functional projection, which I will take to be the inner TopicP in 8.

As for the previous analyses of Type II (or more widely known, the verb-doubling construction), Type II is traditionally analyzed as a LAST RESORT to satisfy both the verb's theta-requirement (i.e. transitive verbs must have an object) and the Mandarin-specific phrase structural constraint in 12 (Huang 1982, Cheng 2007, Tieu 2008, among others).

(12) Phrase Structure Constraint (PSC) (Huang 1982):
Within a given sentence in Chinese, the head (the verb or VP) may branch to the right only once, and only on the lowest level of expansion.

In the last resort analysis, all of the post-main-verbal materials (i.e. DFP/AdvP/RsltP) occupy the complement position to the verb. Since according to the PSC there can only be one complement position to the verb, where the theta-requirement is assumed to be satisfied, the verb copies to create another complement position for the object. However, the following contrast suggests that Type II in the DFP case should not result from the interaction between the PSC and the theta-requirement.

[11]Due to the limit of space, only the examples involving DFPs are demonstrated here, but the same subjectivity applies to Type II in the adverbial/resultative constructions as well.

(13) a. *Lisi <u>nian-de</u> [**zhe-ben shu**]^{DO} [$_{AdvP/RsltP}$ hen kuai/lei]
Lisi read-DE this-CL book very fast/tired
'Lisi read this book very fast.'/'Lisi got tired from reading this book.'

b. Lisi <u>nian-le</u> [**zhe-ben shu**]^{DO} [$_{DFP}$ san tian/ci]
Lisi read-ASP this-CL book three day/time
'Lisi read this book for three days/three times.'

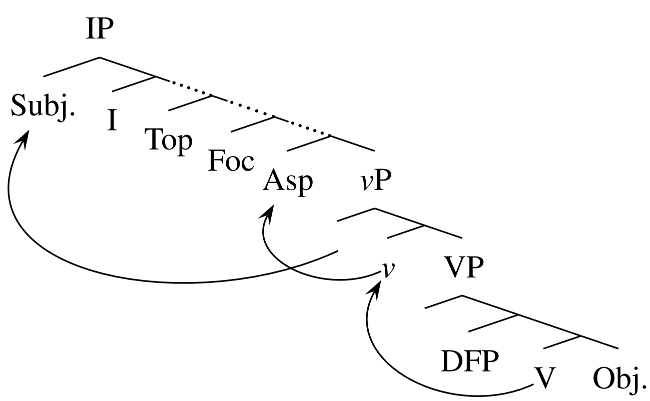
One of the consequences of the last resort analysis is the lack of post-verbal co-occurrence of the object and the other post-verbal phrases, since their competition for the complement position is the cause of the verb-doubling effect. As can be seen in 13, while it is indeed impossible for the object to co-occur with the AdvP/RsltP post-verbally, it is however grammatical for it to do so with the DFP. This would be unexpected under the PSC.

If we maintain the PSC, it would mean that only one of the post-verbal items in 13b can be the syntactic complement to the verb, most likely the object. Therefore, we will treat the DFP as a syntactic adjunct, as opposed to the AdvP/RsltP. It then follows that Type II in the DFP-case should involve a different syntactic derivation than in the other cases (cf. Bartos 2003). Based on their distributional overlap shown in section 2 and the same subjectivity to the functional hierarchy shown in section 3, I will unify Type I and II in the DFP-case in 1 and 2 as SENTENCE-INTERNAL TOPICALIZATION.

4 Analysis

The unified analysis will build on the following structure for Mandarin that involves the functional middle field underlined in 8.

(14)



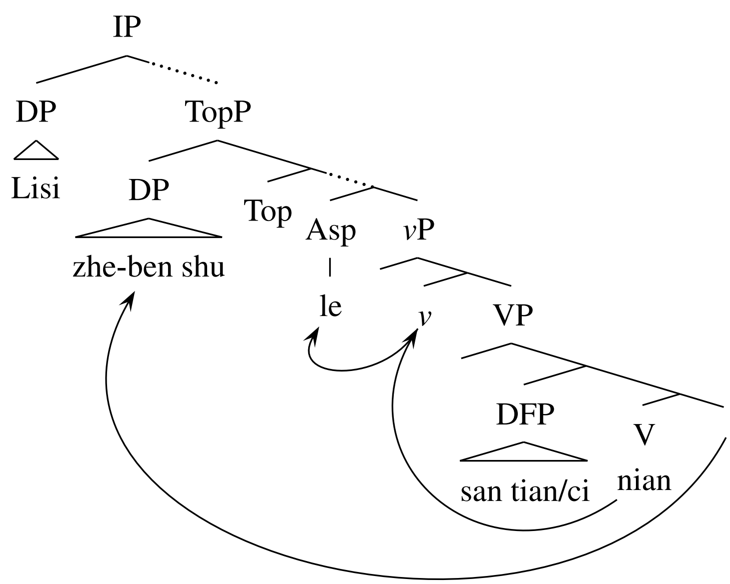
There are some assumptions about this structure: (i) The aspectual morpheme is a syntactic head (Asp^0) in the clausal spine immediately above the *v*P, (ii) the main verb (V) undergoes cyclic head-movement to Asp^0 (Huang, Li, and Li 2009)[12], (iii) the DFP is assumed to be a syntactic adjunct attaching to the intermediate projection inside the VP (Huang, Li, and Li 2009), and (iv) the subject is base-generated in the specifier of the *v*P (Kratzer 1996) and later undergoes argument movement to the specifier of the IP.

4.1 Type I movement

Having laid out the fundamental structure and the associated assumptions, we can now spell out the unified analysis for both Type I and II in the case of DFPs. We will begin with Type I, which is relatively straightforward. For the derivation of Type I, the object undergoes sentence-internal topicalization to the specifier of the inner TopP.

(15) a. Lisi **zhe-ben shu** <u>nian-le</u> san tian/ci (=1)
Lisi this-CL book read-ASP three day/time
'Lisi read this book for three days/three times.'

b.



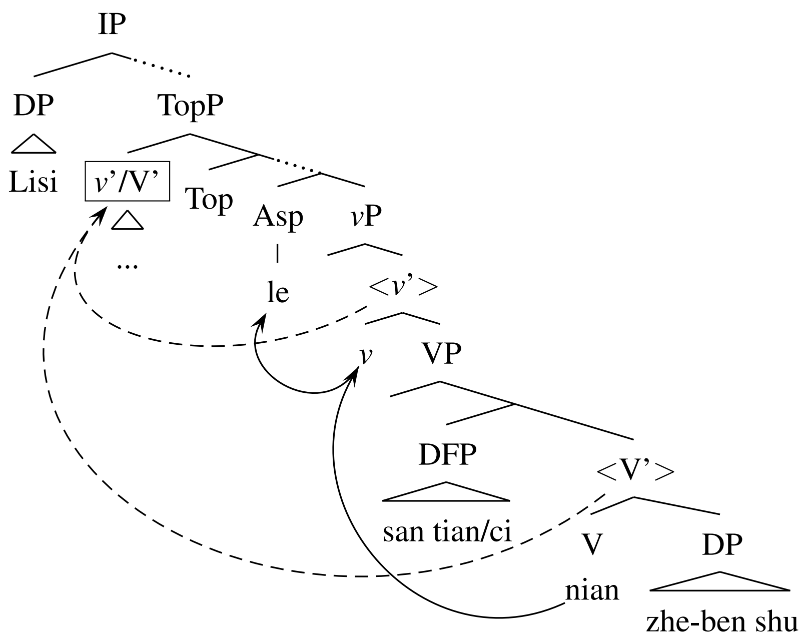
[12] Huang and colleagues (2009) only assume V-to-*v* movement. But since we assume the existence of the aspectual head, we will also assume that the verb moves into Asp^0 for the suffixal needs of Asp^0.

4.2 Type II movement

The derivation of Type II movement is a little more complicated. It involves sentence-internal topicalization, as well as copying of the VP. Unlike in the previous last resort analysis, the trigger of the copy-and-movement of the VP is pragmatic rather than syntactic: VPs can be made conversational topics and become subject to topicalization.

(16) a. Lisi <u>nian</u> **zhe-ben shu** <u>nian-le</u> san tian/ci (=2)
Lisi read this-CL book read-ASP three day/time
'Lisi read this book for three days/three times.'

b.



Above is the detailed derivation process of the Type II example in 2 that we will unpack. Two movement chains are involved, V-movement and VP-topicalization. The former is represented by solid lines, and the latter, dashed lines. Let us focus on the sentence-internal topicalization of the VP. For reasons to be clear later, I propose a COPY-AND-TOPICALIZE mechanism that tracks the head-movement of V by applying to constituents whose left edge V occupies, with an upper bound of *v*P.

This mechanism leads to two possible copy targets given the structure above, i.e. V' and *v*'. Either copy can topicalize, resulting in the displacement of the object to the position preceding the Asp^0-marked main verb. The heads of the two movement chains are pronounced[13], i.e. the higher V and the higher *v*'/V', including the verb head inside the latter due to its non-c-command relationship with the head-moved V (Gouguet 2006). This condition on the pronunciations of the movement copies result in the verb-doubling effect characteristic of Type II. And hypothesizing the *v*P as the upper bound of the copying mechanism accounts for the lack of aspectual marking on the first verb copy, as in 7.

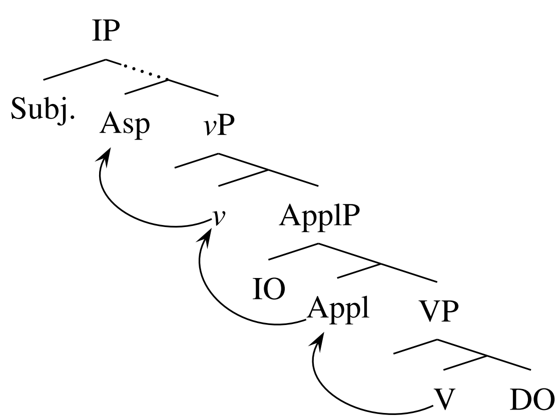
5 Predictions of ditransitive verbs

The proposed copy-and-topicalize mechanism in the previous section predicts various attested word orders for Type II under ditransitive verbs. We will demonstrate the derivations of these word orders with the following example.

(17) Lisi song-le [Mali]^{IO} san ci [liwu]^{DO}
Lisi give-ASP Mary three time present
'Lisi gave Mary presents three times.'

Following Paul and Whitman 2010, we will assume the following structure for ditransitive verbs in Mandarin, where IOs are applied arguments introduced by an APPLICATIVE head ($Appl^0$) (Pylkkänen 2008) right above the VP in the clausal spine.

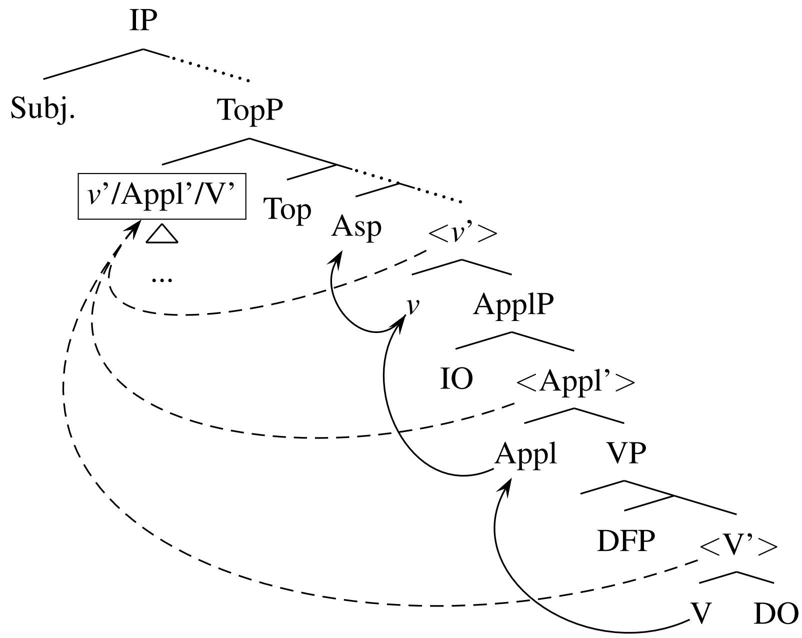
(18)



Applying our analysis for Type II to the ditransitive structure, we get the following result, where our hypothesis that copying tracks V-movement leads to three copy possibilities.

[13]The silent, unpronounced VP copy is indicated by angled brackets as in 16b.

(19)



If what gets copied and topicalized is V' or Appl', then we expect to get the word order where only the direct object (DO) occurs before the main verb, as borne out by 20.

(20) Lisi song [**liwu** $]^{DO}$ song-le [**Mali** $]^{IO}$ san ci
Lisi give present give-ASP Mary three time
'Lisi gave Mary presents three times.'

If what gets copied and topicalized is *v*', then we expect to get the word order where both the DO and IO occur preverbally. Moreover, they are expected to occur in the order where the IO precedes the DO, given the structure in 18.

(21) Lisi song ✔[**Mali** $]^{IO}$ [**liwu** $]^{DO}$ / *[**liwu** $]^{DO}$ [**Mali** $]^{IO}$ song-le san ci
Lisi give Mary present present Mary give-ASP 3 time
'Lisi gave Mary presents three times.'

Finally, the analysis predicts that we will not get a word order where only the IO topicalizes with the verb, since it is not possible to copy a constituent that contains only the IO.

(22) *Lisi song [**Mali** $]^{IO}$ song-le san ci [**liwu** $]^{DO}$
Lisi give Mary give-ASP three time present
'Lisi gave Mary presents three times.'

One might have noted an issue with the copying derivation in 19: How can the DFP stay post-verbal, as in the word orders in 20-21, when the constituent that gets copied and topicalized contains the DFP? We can solve this issue along the lines of Landau 2006, 2007 for the Hebrew verb-copying construction. As Landau (2006, 2007) shows, constituents that occur after the inflected main verb in the Hebrew verb-copying construction are late-merged into the lower unpronounced VP copy. Treating the DFP as being late-merged into the lower VP after the V-movement and Type II movement have taken place ensures the DFP's post-verbal nature at all times.

5.1 More support for Type II (with DFPs) being topicalization

One of the arguments in favor of the aforementioned last resort analysis for Type II is that Type II cannot occur without any of the post-verbal phrases.

(23) *Lisi nian **zhe-ben shu** nian-le
Lisi read this-CL book read-ASP
'Lisi read this book.'

If Type II in the DFP-case is driven by pragmatic force, as we claim to be topicalization, then why does the presence (or absence thereof) of the DFP seem to matter? That is, the syntax of the DFP should play no role in triggering Type II. In fact, I argue that the DFP indeed plays no role and the ungrammaticality of 23 has to do with the notion of topic and comment. I show that once the condition on topic and comment is satisfied, Type II movement is legitimate in the absence of the DFP.

If the copied verbal constituent is indeed made a topic, then the rest of the sentence should be the comment that says SOMETHING ABOUT THE TOPIC. In 23, nothing that follows the copied verbal constituent (now a topic) provides any new information regarding the topic; hence the ungrammaticality. If we can satisfy this condition by having a new-info provider after the topic, then no DFP (or AdvP/RsltP) is needed.

(24) Lisi gan **feiji** gan-shang-le
Lisi catch flight catch-up-ASP
'Lisi caught the flight.'

Gan feiji roughly translates to 'catching flights' in English, where *gan* 'catch' is an open-ended event. However, it can be made a verbal complex with the particle *shang* 'up' that indicates the success of the event. *Shang* provides the new information about the topic of Lisi catching a flight, i.e. it was successful, and renders the DFP unnecessary.

Example 24 also argues against the last resort analysis: Since nothing is occupying the complement position in the main VP, verb-copying should not have happened.

6 Conclusion

Traditionally viewed as independent constructions, bare object preposing (Type I) and verb-doubling (Type II) can in fact achieve a unified analysis as sentence-internal topicalization. It can be done once we segregate the underlying structures of the constructions where Type II is found (DFP-construction vs. manner AdvP/RsltP-construction). With the unified analysis, we can explain the distributional overlap of Type I and II shown in section 2, which would otherwise be accidental. And by implementing a copying mechanism that targets different verbal constituents, we make good predictions about ditransitive word orders under VP-copying. We are also able to account for copying cases where the trigger is not obviously syntactic but pragmatic, as in 24.

References

BADAN, LINDA, 2008. Preposed object and low periphery in Mandarin Chinese. *CISCL Working Papers on Language and Cognition* 2, 19–42.

BARTOS, HUBA, 2003. Mandarin verb-copying is (often) VP-copying. In Paper presented at *the 3rd conference of European Association of Chinese Linguistics*. Belgium: Universiteit Gent.

CHENG, LISA L.-S., 2007. Verb copying in Mandarin Chinese. In *The Copy Theory of Movement*, ed. by Norbert Corver and Jairo Nunes, 151–173. Amsterdam: John Benjamins.

ERNST, THOMAS, and CHENGCHI WANG, 1995. Object preposing in Mandarin Chinese. *Journal of East Asian Languages* 4(3), 235–260.

GOUGUET, JULES, 2006. Adverbials and Mandarin argument structure. In *Empirical Issues in Syntax and Semantics*, ed. by Olivier Bonami and Partricia Cabredo Hofherr, 155–173. Strasbourg: CNRS.

HUANG, JAMES C.-T., 1982. *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

HUANG, JAMES C.-T., 1988. Wo pao de kuai and Chinese phrase structure. *Language* 64, 274–311.

HUANG, JAMES C.-T., AUDREY LI, and YAFEI LI, 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press.

KRATZER, ANGELIKA, 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*, ed. by Rooryck and Zaring, 109–137. Dordrecht: Kluwer.

LANDAU, IDAN, 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9(1), 32–66.

LANDAU, IDAN, 2007. Constraints on partial VP-fronting. *Syntax* 10(2), 127–164.

LI, CHARLES N., and SANDRA A. THOMPSON, 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

PAUL, WALTRAUD, 2002. Sentence-internal topics in Mandarin Chinese: The case of object preposing. *Language and Linguistics* 3(4), 695–714.

PAUL, WALTRAUD, 2005. Low IP and left periphery in Mandarin Chinese. *Recherches Linguistiques de Vincennes* 33, 111–134.

PAUL, WALTRAUD, and JOHN WHITMAN, 2010. Applicative structure and Mandarin ditransitives. In *Argument Structure and Syntactic Relations: A Cross-Linguistic Perspective*, ed. by Maria Duguine, Susana Huidobro, and Nerea Madariaga, 261–282. Amsterdam: John Benjamins.

PYLKKÄNEN, L., 2008. *Introducing Arguments*. Cambridge, MA: MIT Press.

SHYU, SHU-ING, 1995. *The Syntax of Focus and Topic in Mandarin Chinese*. Ph.D. thesis, University of Southern California.

SYBESMA, RINT, 1999. *The Mandarin VP*. Dordrecht: Kluwer.

TIEU, LYN SHAN, 2008. Non-referential verb use in Chinese: A unified verb copying analysis. In *Proceedings of the 20th North American Conference on Chinese Linguistics, vol. 2*, ed. by Marjorie K.M. Chan and Hana Kang, 843–860. Columbus, Ohio: The Ohio State University.

Effects of sensory modality on the interpretation of subjective adjectives: Comparing sight, smell and taste

ELSI KAISER
University of Southern California[*]

1 Introduction

The notion of evaluation is fundamental to our cognition and perception (e.g. Markus & Zajonc 1985), and there exists a range of subjective linguistic expressions, including a class of adjectives known as PREDICATES OF PERSONAL TASTE, such as *fun, tasty, disgusting, amazing*, that reflect evaluative attitudes. To understand these subjective adjectives, one needs to know whose opinion/attitude is being conveyed. Predicates of personal taste are often analyzed as making reference to a judge, attitude holder or evaluator (e.g. Lasersohn 2005, Potts 2007, Stephenson 2007, Patel-Grosz 2012, but see also Pearson 2013). In this paper, I use the term 'attitude holder' (atheoretically) for the individual whose perspective/attitude the adjective is relativized to.

Although predicates of personal taste have received considerable attention in theoretical semantics, to the best of my knowledge, current semantic theories do not make distinctions based on sensory modality (sight, sound, taste, smell, touch). Thus, a sentence such as *It was disgusting* would presumably be analyzed the same way semantically regardless of whether it refers to the taste, smell or visual appearance of a pizza slice, for example. (But see McNally & Stojanovic 2017 on the challenges of aesthetic predicates like *beautiful*.) In this paper, I report two psycholinguistic experiments that investigate whether the identification of the attitude holder of subjective adjectives (specifically, predicates of personal taste) is influenced by the sensory modality that the adjective makes reference to. To see why one might expect sensory modalities to differ, I first review prior work on the biological and social properties of different senses.

1.1 Sensory modalities

It is well-known that the five senses (sight, hearing/audition, taste/gustation, touch/feel and smell/olfaction) are fundamentally different, not only in their biological but also their social-

[*] Many thanks are due to Catherine Wang and Gwenyth Portillo-Wightman for their help with stimulus construction and setting up the studies in Qualtrics. Thanks also to the BLS audience for useful comments and feedback. This material is based upon work supported by the National Science Foundation under Grant No. BCS-1749612.

communicative aspects and general level of subjectivity. In this section, I review the properties of vision, taste and smell, the senses that investigated in the experiments reported here.

Sight is commonly viewed as the dominant sense in most (if not all) human cultures and languages (e.g. San Roque et al. 2015, but contra Aikhenvald & Storch 2013). Biologically, vision is a highly specialized sense in humans, and by some estimates, up to 50% of the cortex is involved in visual functions (Palmer 1999). Research on sensory dominance effects suggests that visual input tends to dominate over auditory input when the two conflict (Colavita 1974, Sinnett et al., 2007, Spence 2009) – in other words, human may have a biologically hardwired preference to rely on the visual modality (but see Aglioti & Pazzaglia 2010).

In addition to these biological factors, the primacy of the visual modality may stem from the fact that it often involves SHARED PERCEPTUAL EXPERIENCES between people. As noted by San Roque et al. (2015), "As a distal sense, it seems likely that sight is one of the most readily and regularly shared perceptual experiences among interlocutors" (p.50). They also note that visual cues are generally viewed as the basic foundation for joint attention (e.g. Moore and Dunham 1995). Furthermore, vision is viewed as providing relatively OBJECTIVE information: Sweetser (1990) claims that vision is "our primary source of objective data about the world" (p.39). Not surprisingly, visual evidence is often considered as more reliable than auditory or other kinds of evidence (e.g. in grammaticalized evidentiality systems).

In contrast, the **olfactory (smell)** and **gustatory (taste) modalities** are regarded as more subjective and variable across people (e.g. Sweetser 1990, Chafe & Nichols 1986, Dubois 2007, Viberg 1984). Caballero and Paradis (2015) note that the Reliability Hierarchy of Evidentiality, which ranks the reliability of sensory experiences, states that "in contrast to the relatively objective and stable nature of visual elements in the world, the perceptions of smell, taste and touch are HIGHLY SUBJECTIVE AND VARIABLE ACROSS HUMAN BEINGS" (e.g. Chafe & Nichols 1986, Viberg 1984). Thus, in contrast to the visual domain (where a person A will tend to assume that she has roughly the same visual experience as person B when they focus their visual attention on the same thing), in the domain of taste or smell A is less likely to assume that she has the same gustatory or olfactory experience as B when they eat or smell the same thing.

1.2 Predicates of personal taste and the importance of experience

Prior theoretical work on predicates of personal taste has largely focused on the question of how to linguistically represent the fact that the meaning of these adjectives is relativized/anchored to the opinion or perspective of an evaluator/judge/attitude holder (e.g., Lasersohn 2005, Stephenson 2007, Patel-Grosz 2012, Pearson 2013). In recent work, researchers have also started to look at how the attitude holder is identified. It is well-known that the speaker is normally the default attitude holder/judge. For example, in 1, the speaker is the preferred attitude holder. However, what about 2, with both a first-person narrator (physically present at the time of the event) and the character in the narrative? Now, the choice of attitude holder is less clear (see Kaiser 2015 for related experimental data):

(1) Speaker says: The muffin was disgusting
Whose opinion is it that the muffin was disgusting? Speaker's OR Addressee's
(2) When I came into the room, Eliza saw the muffin on the platter. It was disgusting.
Whose opinion is it that the muffin was disgusting? The narrator's OR Eliza's

Before taking a closer look at how sensory modality could influence identification of the attitude holder, let us briefly review some of the relevant theoretical work. Recently, it has been argued that predicates of personal taste crucially involve an experiencer argument, in contrast to other kinds of subjective adjectives (e.g. Bylinina 2014; McNally & Stojanovic 2017, see Kaiser & Herron Lee 2017, 2018 for experimental data). In essence, for something to be *fun* or *tasty*, someone must have the relevant experience (usually the speaker). More concretely, as shown in 3 and 4, Bylinina (2014) proposes that with predicates of personal taste, the judge (attitude holder) must be the experiencer of the internal state referred to by the adjective.

(3) JUDGE=EXPERIENCER REQUIREMENT, first take:
A direct statement about someone's internal state can be made only if the judge parameter is set to the same value as the experiencer of this internal state. (Bylinina 2014:58)

(4) $[\![\text{tasty}]\!]^{c;w,t,\mathbf{j}}$ = (i) $\lambda z \lambda x.\ \exists s$ [**taste**(s) & Experiencer(s, z) & Stimulus(s, x) & TASTE(s) $\succ$ d_{st} for j at t in w]; (ii) JUDGE=EXPERIENCER: $j = z$ (Bylinina 2014:52)

1.2.1 Identifying the attitude holder(s) in different modalities

In light of the recent findings regarding the importance of the attitude holder having the relevant kind of experience, let us return to the topic of sensory modality. Given that different modalities involve difference kinds of experiences – and that these experiences can differ in whether they are shared by multiple people or internalized to one person – we might well expect sensory modality to matter for the process of identifying the attitude holder of subjective adjectives. The visual modality often involves shared perceptual experiences between people (e.g. San Roque et al. 2015). Thus, in a context like 5, both the first-person narrator (who enters the room) and Eliza can presumably have the visual experience of seeing the muffin (even though only Eliza is explicitly specified as seeing the muffin). So, when people are asked the *whose opinion* question in 5, both the narrator and Eliza are possible answers (possible attitude holders).

In contrast, the gustatory modality (*taste*) is highly subjective, involves internal experience, and is variable across people (Sweetser 1990). In situation 6, only Eliza tastes the muffin – in all likelihood, the narrator does not have the relevant gustatory experience. Thus, if sensory modality influences how subjective adjectives are interpreted, only Eliza is expected to be available as the attitude holder for 'disgusting' in 6.

(5) When I came into the room, Eliza saw the muffin on the platter. It looked disgusting.
Whose opinion is it that the muffin looked disgusting? The narrator's OR Eliza's
(6) When I came into the room, Eliza tasted the muffin on the platter. It tasted disgusting.
Whose opinion is it that the muffin tasted disgusting? The narrator's OR Eliza's

As regards the olfactory modality, the predictions are less clear. In a situation like 7, Eliza is described as smelling the muffin – but perhaps the narrator (in the same room) can also detect the scent of the muffin. Like vision, smell can occur at a distance and the experience can be shared by multiple people at the same time. Thus, both the narrator or Eliza may be possible attitude holders in 7. However, it is generally assumed that the gustatory modality patterns more like taste in being highly subjective and variable across people (Section 1.1), which may argue against the narrator being available as an attitude holder in 7.

(7) When I came into the room, Eliza smelled the muffin on the platter. It smelled disgusting.
Whose opinion is it that the muffin smelled disgusting? The narrator's OR Eliza's

In sum, while visual input is easily simultaneously experienced by multiple people, taste is more likely to be an internal, individual experience. Additionally, it has been claimed that visual information is treated as more objective and taste-based information as more subjective. Olfactory input seems to lie somewhere in between the shared nature of visual input and the internal nature of gustatory input. The next section describes an experiment that investigated whether people's interpretation of subjective adjectives is influenced by these socio-biological differences between sensory modalities.

2 Experiment 1

Given the striking differences between sensory modalities, two studies were conducted to test whether interpretation of subjective evaluative adjectives (specifically, predicates of personal taste) depends on whether they refer to the visual vs. olfactory (smell) vs. gustatory (taste) domains. Given the sociobiological differences between senses (in particular the shared vs. private nature of the experience) and recent theoretical claims about the importance of the attitude holder being an experiencer, the experiment tested if interpretation of subjective adjectives – specifically, identification of who is the attitude holder – depends on whether they refer to the visual vs. olfactory (smell) vs. gustatory (taste) domains. The experiments reported here focus on vision, smell and taste. They do not investigate hearing/audition and touch, due to challenges associated with incorporating those senses into the within-items experimental design.

2.1 Methods: Participants, design and procedure

Native English speakers (n=56) read two-sentence sequences similar to the examples above and answered questions about them. The study was conducted over the internet and the stimulus items were presented in writing. Participants were told to imagine they were reading extracts from novels, and the term 'narrator' was explained as part of the instructions. The critical sequence of clauses was preceded by a subordinate clause that mentions the speaker/narrator by means of a first-person pronoun (ex.8).[1] This makes available two possible candidates (narrator and character e.g. Eliza) for the *whose opinion* question after each target (ex.9) (Note that the question disambiguates *it* as referring to the muffin/relevant object, not the platter or something else.) This question was presented as a two-alternative force choice. The answers provide a measure of who participants think is the attitude holder of the adjective.

[1] Variants where the preamble mentions the third-person character rather than the first-person narrator, '*When she came into the room*' were also tested, but those are not relevant here as they do not introduce another referent.

(8) a. [sight]

When I came into the room, Eliza saw the muffin on the platter. It looked disgusting.

b. [smell]

When I came into the room, Eliza smelled the muffin on the platter. It smelled disgusting.

c. [taste]

When I came into the room, Eliza tasted the muffin on the platter. It tasted disgusting.

d. [baseline]

When I came into the room, Eliza put the muffin on the platter. It was disgusting.

(9) Whose opinion is it that the muffin {looked/smelled/tasted/was} disgusting?

The narrator's OR Eliza's

The verbs were used to manipulate the senses involved in the item (vision, smell, taste or no sense/baseline).[2] Within an item, the adjective was kept constant. The sense was specified by the verbs in both the first and the last sentences, except for the baseline condition (8d), where it was underspecified. In the baseline, *put* is used to describe the action done by the character in the story, and *is* is used in the second sentence. Thus, no sensory modality is specified. The study included 24 targets (24 different 'vignettes' one of which is shown in 8), which used 12 adjectives (specifically, predicates of personal taste; each used twice), as well as 42 fillers. The items were presented to participants in a Latin-Square design, so that no participant saw more than one version of each item.

2.2 Possible outcomes and their implications

Are there differences between the different sensory conditions in terms of who participants interpret as the attitude holder of the adjective, even though the same adjective is used in all four versions of each item? To the best of my knowledge, current semantic theories of predicates of personal taste do not make any direct predictions about sensory modalities. If we find differences between senses, how can these be captured in theories of evaluativity? One possible avenue is to

[2] Sensory experience often involves multiple senses: As noted by Paradis and Eeg-Olofsson (2013), "we cannot taste something without smelling something, and we cannot taste something without feeling something, and over and above everything is the sight of something" (p.17). However, the verb in the question specifies which sensory modality is being asked about (ex.9) to minimize any ambiguity.

assume an underspecified semantics that leaves room for pragmatic, top-down effects stemming from the social and biological differences between the senses. One such approach has been proposed by Kennedy and Willer (2016:17), who did not look at sensory modalities but who make the point, more generally, that subjectivity is a highly context-sensitive, pragmatic phenomenon that "is not to be explained strictly in terms of any particular semantic parameter, implicit argument, or lexical underspecification" (p.17). This view contrasts with other accounts of how to encode the attitude holder of predicates of personal taste (e.g. Lasersohn 2005, who proposes a judge parameter, Bylinina 2014 who argues in favor of implicit arguments), but would allow us to explain potential sensory modality effects without having to complicate the lexical entries of the adjectives themselves.

In addition to the question of WHETHER differences exist, this work also explores WHAT KINDS OF DIFFERENCES we might find: In 8a and 8c, Eliza is the subject of *saw* and *tasted*. Given that gustatory experiences in general involve a person's internal subjective experience and are variable across individuals*,* the prediction is that Eliza will be interpreted as the attitude holder of *disgusting* in 8c, with *taste*. However, as visual experiences often involve shared perceptual experiences and tend to be more stable/consistent across individuals, the first-person narrator may also be interpreted as the attitude holder for 'disgusting' in 8a, with *see*. Thus, if the attitude-holder identification process with subjective adjectives is sensitive to the sensory information on the verb, there may be more narrator responses (and less character responses) with *see* than *taste*. The predictions for *smell* are unclear: It involves more shared perceptual experiences than *taste* but is intuitively less constant across individuals than *see*.

3 Experiment 1 results and discussion

The proportion of 'character's opinion' responses are shown in Figure 1. (The proportion of 'narrator's opinion' responses is the inverse of the 'character opinion' responses, as the study used a two-alternative forced-choice task). The results were analyzed used logistic mixed-effects regression models (lmer), using R (R Core Team, 2018) as they are better suited for this kind of categorical data than ANOVAs.

The baseline condition (no sensory modality specified) elicited mostly 'narrator's opinion' responses and less than 25% 'character's opinion' responses. This is in line with existing claims from the theoretical literature that the speaker (or writer) is the default attitude holder: In contexts where no sensory modality is specified and the subjective adjective occurs in a sentence with the copular *is*, participants overwhelmingly interpret the first-person narrator as the attitude holder. Indeed, the proportion of character's opinion responses is significantly lower than chance ($\beta = -1.415$, $SE = 0.0298$, $z = -4.739$, $p < 0.0001$). (From-chance analyses were conducted using intercept-only logistic regression models.)

Crucially, however, the default preference to interpret the speaker/writer as the attitude holder of the subjective adjective vanishes in the other three conditions. Once the character in the narrative is described as the subject of sensory verb – whether it is seeing, smelling or tasting – then that character becomes the preferred attitude holder for the subjective adjective. As can be clearly seen in Figure 1, the other three conditions differ strikingly from the baseline. Regardless of which sensory modality is specified, all three conditions elicit a higher-than chance rate of character's opinion responses (taste: $\beta = 2.881$, $SE = 0.764$, $z = 3.769$, $p < 0.001$, smell: $\beta = 1.808$, $SE = 0.393$, $z = 4.601$, $p < 0.0001$, see: $\beta = 0.8699$, $SE = 0.328$, $z = 2.653$, $p < 0.01$).

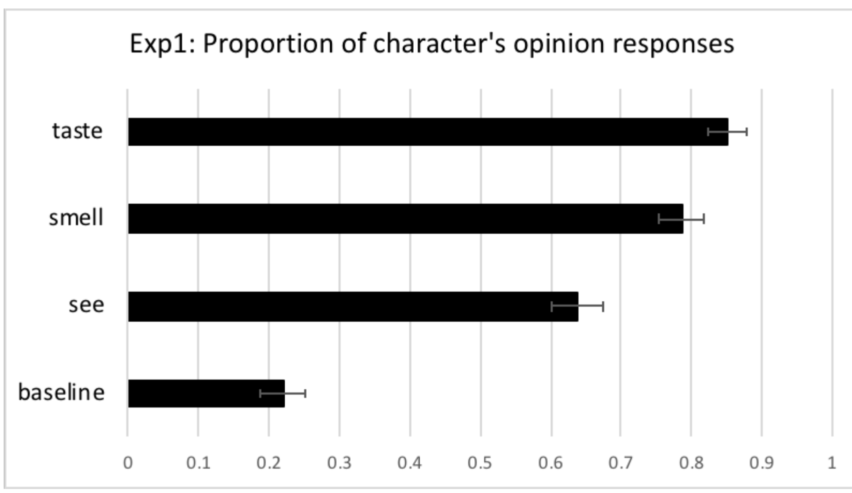


Figure 1. Proportion of character's opinion responses in Experiment 1. Error bars show +/- 1 SE. (The proportion of narrator opinion responses is the inverse of character responses.)

Furthermore, when the conditions are compared to each other, we find that the rate of character opinion responses is higher (and the rate of narrator responses lower) in the *smell* and *taste* conditions than the *see* conditions ($p < .003$ and $p < .001$ respectively) or the baseline condition ($p < .001$ and $p < .001$ respectively). (As expected based on Figure 1, the taste and smell conditions do not differ significantly from each other.) Thus, although all three sensory conditions show a preference to interpret the character as the attitude holder (rather than the narrator), this preference is significantly stronger with *taste* and *smell* than with *see*.

In sum, sensory modality significantly impacts the process of identifying the attitude holder of predicates of personal taste. Contexts involving the gustatory and olfactory modalities elicit more character's opinion responses than contexts involving the visual modality. It is important to acknowledge that the current work was not designed to definitively answer the question of WHY the differences are the way they are – the main aim was to see if differences between modalities EXIST. However, the finding that contexts involving vision elicit significantly fewer character responses (more narrator responses) is in line with observations concerning the shared-experience nature of vision: Not only the character, but also the narrator can be interpreted as the attitude holder of the subjective adjective, as both are receiving and experiencing visual input. The fact that smell and taste pattern alike (and elicit more character responses) is in line with prior claims that these modalities tend to involve more internal (and more subjective) experiences than vision.

4 Experiment 2

Experiment 2 builds on Experiment 1, and has two main aims: One aim of this study is to see if linguistic cues (unrelated to sensory modality) can be used to modulate the availability of the speaker/writer as an attitude holder. We saw in Experiment 1 that although the speaker/writer is the preferred attitude holder in the baseline condition (no sensory modality mentioned), once the narrative contains a character who experiences sensory input, the speaker/writer becomes significantly less likely to be construed as the attitude holder. Experiment 2 uses intensifiers (e.g. *totally, absolutely*) to modify adjectives (e.g. *totally disgusting*) to see if this can boost the availability of the speaker/writer as the attitude holder.

The choice of intensifiers is motivated by observations in Beltrama (2018), who discusses uses of *totally* with predicates that do not express a bounded scale. In addition to the traditional examples with predicates that refer to upper-bounded scales like *full* in 10a, *totally* can also be used with adjectives like *awesome* (ex.10b) which do not lexicalize a bounded scale (examples from Beltrama). In these kinds of contexts, "the use of totally contributes to strengthening the speaker's commitment towards the utterance" (Beltrama 2018). In other words, by combining *totally* with *awesome* – or other open-scale adjectives like *tasty* or *disgusting* – a speaker can signal the strength of their belief that the proposition should be added to the Common Ground.

(10) a. The bus is totally full.
b. Skiing around Salt Lake is totally awesome.

Thus, one could perhaps hypothesize that use of a linguistic expression that strengthens the speaker's commitment towards the utterance would render the speaker more available as an attitude holder, in a context where the speaker (writer) and a character in the narrative are both potentially available as competing attitude holders. This is the idea tested in Experiment 2.

It is important to note that Beltrama focuses on *totally*, which – as he shows – does not behave in the same was as *really*, for example (see Romero and Han 2004 on *really*). Experiment 2 tested a range of intensifiers, to avoid excessive lexical repetition within the experiment. I acknowledge that grouping together multiple intensifiers is likely to be an oversimplification and that further work is needed to better understand the differences between various intensifiers.

The second main aim of Experiment 2 is to see if the basic outcome of Experiment 1 can be replicated with a new group of participants (and stimuli with intensifiers). In recent years, the notion of replicability has gained increasing visibility in psychological and psycholinguistic research. Especially when a new research area is tested experimentally (as is the case here), including some amount of replication is helpful for establishing the credibility of the results.

4.1 Methods: Participants, design and procedure

The design, procedure and methods were the same as Experiment 1, except that in all target items, the subjective adjective in the final clause was preceded by an intensifier (e.g. *totally, absolutely, really, extremely)*, as shown in 11. I report data from 56 native English speakers, none of whom had participated in Experiment 1.

(11) a. [sight] When I came into the room, Eliza saw the muffin on the platter. It looked really disgusting.
b. [smell] When I came into the room, Eliza smelled the muffin on the platter. It smelled really disgusting.

c. [taste] When I came into the room, Eliza tasted the muffin on the platter. It tasted really disgusting.

d. [baseline] When I came into the room, Eliza put the muffin on the platter. It was really disgusting.

(12) Whose opinion is it that the muffin {looked/smelled/tasted/was} really disgusting?
The narrator's OR Eliza's

4.2 Possible outcomes

One set of predictions for Experiment 2 parallels those for Experiment 1: If the attitude-holder identification process with evaluative adjectives is sensitive to the sensory information on the verb, we expect to see differences between the sensory modalities – more specifically, we expect to more character responses with *taste* and *smell* than *see*, in line with Experiment 1.

A second set of predictions concerns the potential effects of intensifiers. If intensifiers make the speaker/writer of the sentence (in our case the first-person narrator) more available as an attitude holder, then we expect to see more narrator responses in Experiment 2 than in Experiment 1, at least in contexts where a shared perceptual experience is possible. In other words, in the visual condition – and perhaps in the olfactory condition – we expect to see more 'narrator's opinion' responses in Experiment 2 than in Experiment 1. This is because although the character in the story is described as seeing/smelling the object being described, the first-person narrator is also present in the same space and thus can also experience the gustatory or olfactory sensory input. In the taste condition, which involves an internal experience by the character mentioned in the narrative, it is unlikely that boosting the availability of the narrator will have an effect, as the narrator is not described as being involved in the tasting event.

5 Experiment 2 results and discussion

The proportion of trials on which participants answered that the subjective adjective reflects the opinion of the character in the story (rather than the narrator) is shown in Figure 2. As in Experiment 1, the proportion of 'narrator's opinion' responses is the inverse of the 'character's opinion' responses (due to the two-alternative forced-choice design). Similar to Experiment 1, the baseline condition elicits a low proportion of 'character's opinion' responses (significantly below chance: $\beta = -2.05$, $SE = 0.55$, $z = -3.73$, $p < .001$). In line with what we saw in Experiment

1, other things being equal, subjective adjectives tend to be interpreted as anchored to the speaker or writer of the sentence (here, the first-person narrator).

The conditions involving the gustatory modality and the olfactory modality elicit a higher-than-chance rate of 'character's opinion' responses (taste: $\beta = 2.235$, $SE = 0.553$, $z = 4.04$, $p < .0001$), smell: $\beta = 1.122$, $SE = 0.382$, $z = 2.935$, $p <. 001$), in line with Experiment 1. Thus, as we already saw in Experiment 1, when the subjective adjective describes a taste or smell that the character experiences, the character is interpreted as the attitude holder (though smells could in principle be a shared experience between multiple people). The condition involving the visual modality results in an at-chance rate of 'character's opinion' responses and 'narrator's opinion' responses ($\beta = 0.016$, $SE = 0.33$, $z = 0.048$, $p > .96$).

Indeed, similar to what we saw in Experiment 1, when we compare the conditions to each other, we find that the rate of character opinion responses is higher (and the rate of narrator responses lower) in the *smell* and *taste* conditions than the *see* condition ($p < .001$ and $p < .001$ respectively) or the baseline condition ($p < .001$ and $p < .001$ respectively). So far, the results for Experiment 2 largely replicate Experiment 1, indicating that (i) participants' interpretation of who is the attitude holder of the subjective adjective depends significantly on the sensory modality and (ii) the gustatory and olfactory modalities have the strongest preference for shifting away from the narrator/writer/speaker to the character in the story. (We discuss the *smell* vs. *taste* comparison below.)

What about potential effects of intensification? The presence of intensifiers was predicted to increase the availability of the narrator as an attitude holder, especially in contexts where shared perceptual experiences are possible, i.e., with *see* and maybe *smell*. Indeed, a comparison of Figures 1 and 2 shows that the rate of narrator responses with *smell* and *see* is higher in Experiment 2 than Experiment 1 (significantly higher with *smell*: $p < .05$, marginally higher with *see*: $p = .052$). There are no significant differences between Experiments 1 and 2 in the baseline condition or the *taste* condition. The differences between Experiments 1 and 2 suggest that intensification can indeed boost the likelihood of the first-person narrator being interpreted as the attitude holder in exactly those modalities where shared experience is possible, i.e., the speaker/writer/narrator can also be an experiencer. Thus, the process of identifying the attitude holder is influenced by multiple constraints, including effects of sensory modality as well as the salience/availability of the first-person narrator.

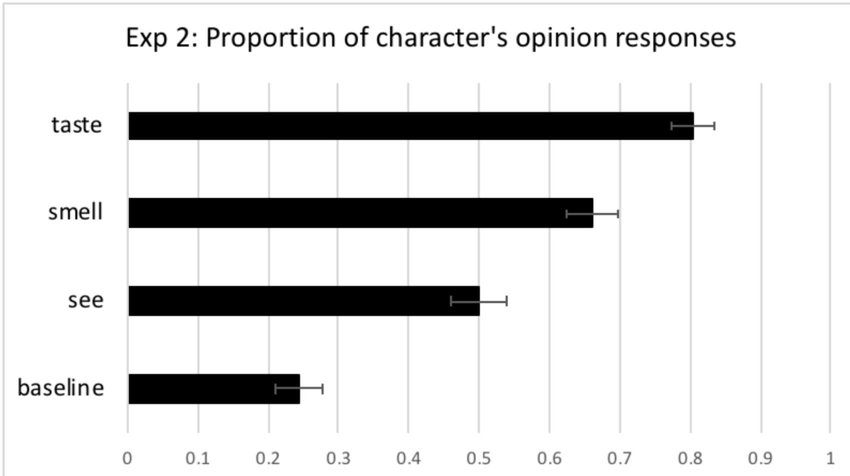


Figure 2. Proportion of character’s opinion responses in Experiment 2. Error bars show +/- 1 SE.

Potential further evidence for the effects of intensification comes from the finding that (i) though the proportion of character responses in the *smell* and *taste* conditions do not differ in Experiment 1, (ii) *taste* elicits more character responses in Experiment 2 than *smell* ($p < .004$). This difference fits with the finding that in Experiment 2, the proportion of character responses with *smell* is lower than in Experiment 1, since the narrator is more likely to be considered as a potential attitude holder in Experiment 2. This decrease of character responses in Experiment 2 does not occur with *taste*, presumably because taste is an internal, non-shared experience. As a result, a difference emerges between *taste* and *smell* in Experiment 2, further confirming that identification of the attitude holder is sensitive to multiple constraints.

6 General discussion

Whereas most prior work on subjective linguistic expressions has focused the question of how to linguistically represent and encode an attitude holder/judge/evaluator for subjective expressions, this paper reports two experiments on how comprehenders identify the attitude holder when multiple candidates are (in principle) available. Specifically, the experiments tested whether, in the case of predicates of personal taste, identification of the attitude holder is modulated by the sensory modality that the situation makes reference to. The studies tested sight, taste and smell, which differ in the social and biological properties of the relevant sensory experience.

The results show that differences in sensory modality significantly impact the process of identifying the attitude holder of subjective adjectives. Participants are more likely to interpret the first-person narrator as being the attitude holder with *see* when compared to *taste* and *smell*, and conversely are more likely to interpret a character in the narrative as being the attitude holder with *taste* and *smell*, relative to *see*. I tentatively suggest that these findings are likely attributable to the fact that taste and smell are largely internal experiences and vary across individuals (and thus the only plausible attitude holder is the one explicitly described as having the experience), whereas seeing something is a perceptual experience often shared by multiple individuals at the same time (and thus the first-person narrator can also possibly share the experience). Further work is needed to assess these ideas in more detail. As a whole, this work shows that when investigating attitude holders accessible to subjective expressions, one needs to pay attention to the sensory modality involved in the experience.

Furthermore, Experiment 2 found that intensifiers (e.g. *absolutely disgusting*) also influence comprehenders' identification of the attitude holder: In contexts where the perceptual experience can be shared, presence of an intensifier boosts the rate of first-person narrator interpretations. Thus, attitude holder identification is a process guided by multiple constraints, rooted in different linguistic and cognitive sources.

Although theories of predicates of personal taste do not explicitly compare sensory modalities (to the best of my knowledge), they are not incompatible with sensory-modality effects. I view the present results as compatible with claims that the attitude holder (of predicates of personal taste) must be an experiencer (e.g. Bylinina 2014, McNally & Stojanovic 2017). Broadly speaking, if we treat subjectivity as context-dependent (e.g. Kennedy & Willer 2016), we can derive the sense-based differences from the biological and social properties of sight, taste and smell, without needing to complicate the lexical entries of individual subjective adjectives.

REFERENCES

AGLIOTI, SALVATORE MARIA, AND MARIELLA PAZZAGLIA, MARIELLA. 2010. Sounds and scents in (social) action. *Trends in Cognitive Sciences* 15(2).47-55.

AIKHENVALD, ALEXANDRA Y., AND ANNE STORCH. 2013. Linguistic expression of perception and cognition: A typological glimpse. In *Perception and cognition in language and culture*, ed. by Alexandra Y. Aikhenvald & Anne Storch (eds.), 1-46. Leiden: Brill

BELTRAMA, ANDREA. 2018. *Totally* Between Subjectivity and Discourse. Exploring the Pragmatic Side of Intensification. *Journal of Semantics*, 35(2).219-261.

BYLININA, LISA. 2014. *The Grammar of Standards.* LOT Dissertation Series.

CABALLERO, ROSARIO, AND CARITA PARADIS. 2015. Making sense of sensory perceptions across languages and cultures. *Functions of Language,* 22(1).1-19.

CHAFE, WALLACE L. AND JOHANNA NICHOLS. (eds.). 1986. *Evidentiality: The linguistic coding of epistemology*. Norwood, NJ: Ablex

COLAVITA, FRANCIS. 1974. Human sensory dominance. *Perception & Psychophysics*, 16.409-12.

DUBOIS, DANIÈLE. 2007. From psychophysics to semiophysics: Categories as acts of meaning. A case study from olfaction and audition back to colors. In *Speaking of colors and odors,* ed. by Martina Plümacher and Peter Holz, 167–185. Amsterdam: Benjamins.

HARRIS, JESSE. 2012. *Processing Perspectives.* Dissertation, UMass.

KAISER, ELSI. 2015. Perspective-shifting and free indirect discourse: Experimental investigations. *Semantics and Linguistic Theory 25.*346-372.

KAISER, ELSI, AND HERRON LEE, JAMIE. 2017. Experience matters: A psycholinguistic investigation of predicates of personal taste. *Semantics and Linguistic Theory* 27.323-339

KAISER, ELSI, AND HERRON LEE, JAMIE. 2018. Predicates of personal taste and multidimensional adjectives: An experimental investigation. *Proceedings of the 35th West Coast Conference on Formal Linguistics (WCCFL35*), 224-231.

KENNEDY, CHRISTOPHER, AND MALTE WILLER. 2016. Subjective attitudes and counterstance contingency. *Semantics and Linguistic Theory* 26.913-933.

LASERSOHN, PETER. 2005. Context dependence, disagreement, and predicates of personal taste'. *Linguistics and Philosophy* 28.643-686

MARKUS, HAZEL, AND ROBERT B. ZAJONC. 1985. The cognitive perspective in social psychology. In *The handbook of social psychology (*3rd ed.), ed. by Gardgner Lindzey & Elliot. Aronson, 137-230. New York: Random House.

MCNALLY, LOUISE, AND ISIDORA STOJANOVIC. 2017. Aesthetic adjectives. In *The Semantics of Aesthetic Judgments*, ed. by James Young, 17-37. Oxford: Oxford University Press.

MOORE, CHRIS, AND PHILIP J. DUNHAM (eds.). 1995. *Joint attention: Its origins and role in development.* Hillsdale, NJ: Erlbaum

PALMER, STEPHEN E. 1999. *Vision science: Photons to phenomenology.* Cambridge: Bradford.

PARADIS, CARITA, AND MATS EEG-OLOFSSON. 2013. Describing sensory experience: The genre of wine reviews. In *Metaphor in and across genres. Thematic issue of Metaphor and* Symbol, ed. by Rosario Caballero, 28(1). 22–40

PATEL-GROSZ, PRITTY. 2012. *(anti-)Locality at the Interfaces.* PhD dissertation, MIT.

PEARSON, HAZEL. 2013. A judge-free semantics for predicates of personal taste. *Journal of Semantics* 30(1).103-154.

POTTS, CHRISTOPHER. 2007. The expressive dimension. *Theoretical Linguistics* 33 (2).165-98.

R CORE TEAM. 2018. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

ROMERO, MARIBEL, AND CHUNG-HYE HAN. 2004. On yes/no questions. *Linguistics and Philosophy* 27.609–658.

SAN ROQUE, LILA ET AL. 2015. Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics* 26(1).31-60

SINNETT, SCOTT; CHARLES SPENCE; AND SALVADOR SOTO-FARACO. 2007. Visual dominance and attention: The Colavita effect revisited. *Perception & Psychophysics,* 69.673–686.

SPENCE, CHARLES. 2009. Explaining the Colavita visual dominance effect. *Progress in Brain Research* 176.245-258.

SWEETSER, EVE. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure.* Cambridge: Cambridge University Press

VIBERG, ÅKE. 1984. The verbs of perception: A typological study. *Linguistics* 21(1).23-62.

Gradable Ideophones, Scales, and Maximality in Grammar

KOJI KAWAHARA
Nagoya University of Foreign Studies[*]

1 Introduction

Japanese is known to have a rich system of ideophones that are a class of referential words, evoking a vivid, sensational feeling, or depictive meanings (Kita 1997; Tamori & Schourup 1999; Akita 2009; Dingemanse 2011; Dingemanse 2012; Dingemanse 2015; Dingemanse & Akita 2017; Dingemanse 2017). Native speakers of Japanese have the intuition that ideophones induce direct, sensory impressions, allowing them to detect the non-arbitrary or iconic relations between sounds and meanings. The focus of this paper is on the gradability of ideophones, in which the core meaning of gradability is based on the abstract representation of measurement or scales. Based on degree constructions, I will show that at least some ideophones are symbolic, in that they will readily participate in grammatical formations. A consequence of the present study is that some ideophones can be analyzed in terms of the ontology of degrees and that their sensational flavor is due to the flexibility of a standard of comparison. Other ideophones are iconic and not gradable; their grammatical formation is limited in some way.

The organization of this paper is as follows. Section 2 discusses the semantic properties and curiosities of ideophones. Based on the discussion in section 2, section 3 introduces gradable ideophones and provides a formal semantic analysis of gradable ideophones. Section 4 concludes.

[*] *I thank audiences at Nagoya Gakuin University, Osaka University and BLS 44 for helpful comments and discussions. I am particularly indebted to Kimi Akita and Mark Dingemanse for detailed and thoughtful comments. Of course, all remaining shortcomings and incoherencies are my responsibility. This paper is based upon work supported by MEXT/ JSPS Grant-in-Aid for Early-Career Scientists Grant Number 18K12385.*

2 Ideophones

2.1 Iconic vs. Symbolic

Saussurean theory of language has been influential in the contemporary study of linguistic signs. In Saussurean semiotics, the relation between signifié and signifiant is arbitrary, whereby concepts and phonological signs are not necessarily similar but are motivated by social convention.

Another influential semiotics is Peirce's so-called "Theory of Signs" (Burch 2017), in which SYMBOL and ICON are defined as follows:

(1) a. SYMBOL: A mode in which the signifier does not resemble the signified object. The relation between them is fundamentally arbitrary or conventional (e.g., language, numbers, morse code, and traffic lights.)

b. ICON: A mode in which the signifier is conceived as resembling or imitating the signified object. The relation between them is somehow similar (e.g., portraits, cartoons, and imitative gestures.)

Ideophones are special in that they are both symbolic and iconic. In the next subsection, I show several important properties of ideophones.

2.2 Sound-Symbolic Words

Ideophones are sound-symbolic words and are sometimes called ONOMATOPOEIA or MIMETICS. While the term ONOMATOPOEIA is most generally used for sound-symbolic words in Japanese linguistics (Kakehi & Tamori 1993; Tamori & Schourup 1999), it is usually understood to be limited to words of imitative sound (Moore 2015; Dingemanse 2018). Hence, it is deemed a subset of ideophones.

MIMETICS is a translation from Sino-Japanese/Korean terms for subtypes of sound-symbolic words; they can be divided into the following four types (Akita 2009):[1]

[1]Kindaichi (1978) provides the five classes of mimetics–GISEIGO, GIONGO, GITAIGO, GIYOOGO, AND GIZYOOGO–but the classification is not important here.

(2) a. GISEIGO (phonomime): Ideophone/onomatopoeia that mimics a human or animal vocalization. (e.g., waNwaN (a dog's barking) or nyaanyaa (a cat's miaow).)

b. GIONGO (phonomime): Ideophone/onomatopoeia that mimics natural noises by inanimate objects (e.g., doNdoN (hitting something) or gorogoro (thunder).)

c. GITAIGO (non-phonomime/ phenomime): Ideophones that describe states or motion (e.g., kaNkaN (the sun blazing) or tekipaki (briskly, actively).)

d. GIZYOOGO (non-phonomime/ phenomime): Ideophones that describe psychological states (e.g., yakimoki (anxiously) or bikubiku (scared))

According to Doke (1935:118), an IDEOPHONE is "a vivid representation of an idea in sound. A word, often onomatopoetic, which describes a predicate, qualificative or adverb in respect to manner, color, sound, smell, action, state or intensity." Since the term ideophone is the most flexible one, I use it in this paper.

According to Dingemanse, IDEOPHONES are marked words that depict sensory imagery. He characterizes the concept as follows.

(3) a. Ideophones are marked structurally, morphologically and phonologically.

b. Ideophones are words, conventionalized items with specifiable meanings.

c. Ideophones are depictions that are special in the way they signify their referents (Dingemanse 2011; Dingemanse 2017).

d. Ideophones depict sensory imagery: perceptual knowledge that derives from sensory perception of the environment and the body (Dingemanse 2011; Dingemanse 2012).

The distinction between 'description' and 'depiction' is important for Dingemanse. To illustrate this distinction, "walking with a limp" is a description and the ideophone tjáḍitjaḍi (Ewe (Westermann 1907)) is a corresponding depiction. Under Dingemanse's analysis, 'description' is an arbitrary sign, interpreted according to a conventional symbol system, while 'depiction' is a performance that invites us to pretend that we are actually experiencing the scene depicted.[2] The dichotomy of 'depiction' and 'description' is summarized in Table 1.

	Depiction	**Description**
Form-meaning mapping	iconic	arbitrary
Building blocks	gradient markings	discrete symbols
Interpretation	imagine to interpret	decode to interpret

Table 1: 'Depiction' vs. 'Description' (Dingemanse 2015)

Bolinger (1968:17) claims that "Language is digital, not analog: its units function by being either present or absent, not by being present in varying degrees." Bolinger's perspective is applied to DESCRIPTION, but if ideophones are depictive, it is necessary to admit that at least some expressions are analog, whereby graded and dividable phenomena are possible.

The frequency and availability of ideophones are subject to cross-linguistic variations. Some languages have a rich system of ideophones, and some do not. Even among rich systems of ideophones, the systems show several variations of syntactic categories. Based on Siwu, a Niger-Congo language spoken in Ghana, Dingemanse (2017) classifies the usage of ideophones as shown in Table 2.

Construction	Tokens	Proportion
Adverbial	101	46%
Complement	56	26%
Holophrase	27	12%
Adjectival	13	6%
Predicative	11	5%
Other	11	5%
Total	219	100%

Table 2: Ideophone constructions in the Siwu Corpus (Dingemanse 2017)

The distribution of Siwu ideophones is very similar to Japanese, as has been pointed

[2]Kita (1997), citing Diffloth (1972), associates the depictive nature of ideophones with imagistic and affective one and Güldemann (2008) treats the depictive nature within a domain of MIMETICS.

out by Tamori & Schourup (1999); Akita (2009). As shown in (4), adverbial ideophones are quite abundant in Japanese. The ideophones modify the respective verbs.

(4) a. Densya-ga gooo-to hasiri satta.[3]
train-NOM IDPH-QUOT run left
'The train left gooo [roaringly].'

b. Ame-ga zaazaa(-to) futta.
rain-NOM IDPH-QUOT fell
'It rained zaazaa [pouringly].'

The ideophones below are the complements of the verbs.

(5) a. Mimi-ga gorogoro(-to) kikoeru.
ear-NOM IDPH(-QUOT) sound
'The ears sound gorogoro [rumbling].'

b. Onaka-ga kirikiri(-to) kanziru.
stomach-NOM IDPH(-QUOT) seem
'My stomach seems kirikiri [stabbing].'

Holophrastic usage is possible, in which the ideophones appear by themselves.[4]

(6) a. Dobaaa!
IDPH
'Dobaaa [forcefully] (some liquid falls out).'

b. Dooon!
IDPH
'Dooon [kaboom] (something falls down).'

[3] Ideophones in Japanese are frequently accompanied by to, which functions as a quotative marker. According to Tamori & Schourup (1999), the following patterns are observable. As for the property of to, consult Akita & Usuki (2016).

i. to is optional (CVCV-CVCV): katakata(-to) (clattering), zabuzabu(-to) (splashing)
ii. to is not attached (degree adverbial, frequency): tyokutyoku (often), sukkari (completely)
iii. to is preferable (VCVri, CVCVN): fuwarifuwari-to (softly), pikaripikari-to (twinkling)
iv. to is obligatory: pat-to (suddenly), pon-to (popping)

[4] I thank Kimi Akita for these examples.

Adjectival uses are possible if the ideophones are followed by no.

(7) a. Pikapika-no kutu-o katta.
IDPH-COP shoes-ACC bought
'I bought pikapika [brand-new] shoes.'

b. Kotikoti-no koori-ga suki da.
IDPH-COP ice-NOM like be
'I like kotikoti [stiffened] ice.'

Predicative uses are possible if the ideophones are followed by a light verb.

(8) a. Asimoto-ga pokapoka-da.
foot-NOM IDPH-be
'My feet are pokapoka [warm].'

b. Atama-ga gangan-sita.
head-NOM IDPH-be
'My head was gangan [splitting].'

In the following section, I will show that some ideophones are gradable, symbolic signs that are associated with scales.

3 Gradable Ideophones

Gradability is typically associated with adjectives and adverbs. Since many adjectives and adverbs denote comparisons among objects based on some standard, they are available in comparative constructions, in degree constructions accompanied by too or so, and are combinable with measure phrases. Typically, gradable adjectives and adverbs are inflected in comparative constructions (e.g., -er or more); thus, the notion of gradability is highly connected with grammatical terms that are supported by morphological evidence. In the case of Japanese, comparative morphemes are unavailable, but gradability is typically expressed by adjectives and adverbs. The purpose of this section is to investigate the gradability of ideophones and to show that gradable ideophones are based on an abstract representation of measurement or scales.

3.1 Gradability

Gradability can be confirmed by the availability of degree constructions, typically comparisons. Japanese comparisons are made possible by the standard marker yori(mo), which expresses a standard of comparison for its counterpart or target of comparison (Sawada 2013). The non-phonomime ideophones, or GITAIGOs, can host comparisons and are compatible with degree modifiers (e.g., totemo 'very'). The ideophone dotabata expresses the state of making noises and doing something hard. The comparison in (9a) is, for example, dependent on the standard of dotabata–that Bella ran more noisily and harder than Alfie. Hence, the non-phonomime ideophones can provide scales.[5]

(9) a. Bella-ga Alfie yori(mo) dotabata(-to) hasitta.
Bella-NOM Alfie than IDPH ran
'Bella ran more dotabata [noisily and hard] than Alfie.'

b. Bella-ga totemo dotabata(-to) hasitta.
Bella-NOM very IDPH ran
'Bella ran very dotabata.'

GIONGOs, phonomime ideophones or onomatopoeia, are interesting in that metaphorical interpretation or extent meaning is somehow required. DoNdoN in (10) is originally derived from the sound of hitting. The standard seems to be derived from the ideophone, expressing the extent of hitting in the examples in (10) where Holms hits the door hard. Although the ideophone is based on the sound of hitting, the standard is not necessarily relevant to the sound of hitting: it is agnostic about the extent of the sound of hitting. In fact, the extent reading is possible even in the absence of ideophones, and an important generalization about Japanese degree constructions is that the extent reading is possible even in the absence of overt degree expressions.

[5] Of course, not all ideophones are gradable. Some of them are summarized below:

i. One time: fut(Q)o (by accident), saQ (suddenly)

ii. Phonomime: kaN (clang), baN (bang)

iii. Non-phonomime: saaQ (promptly), gaaN (bummer)

Another reading is also possible in the reduplicated form of ideophones. DoNdoN is the dual form of don and is regarded as the plural form of ideophones, in which the hitting event by Holms does not necessarily correspond to two: the ideophone is usable as long as the hitting event is more than one.[6] The standard can also be based on the number of hitting events, and thus, the comparison based on the number of hitting events is possible in (10a). Again, in the absence of the ideophone, the number reading is possible. Hence, another generalization is that the number reading is possible if the number of events is repeatable and thus countable or semelfactive.[7]

(10) a. Holms-ga Watson yori(mo) doa-o (doNdoN) tataita.
Holms-NOM Watson than door-ACC IDPH hit
'Holms hit the door more doNdoN [banging] than Watson.'

b. Holms-ga doa-o totemo (doNdoN) tataita.
Holms-NOM door-ACC very IDPH hit
'Holms hit the door very doNdoN [banging].'

Another type of phonomime or onomatopoetic word, based on the sound of animals (including humans), is GISEIGO. This type of ideophone also receives metaphorical or extent readings in degree constructions. WaNwaN describes the sound of barking, akin to bow-wow. Degree constructions are possible with this ideophone, and it denotes the extent of barking, where Bella barked hard in the following examples. Again, the extent reading is possible in the absence of the ideophone. Since waNwaN is a reduplicated form and thus is plural, the standard can also be based on the number of barking events in

[6] The iconicity of events and ideophones is found in a single form. In the following example, the hitting event occurs only once, and thus, degree construction based on doN is not possible.

i. Holms-ga doa-o doN-to tataita.
Holms-NOM door-ACC IDPH hit
'Holms hit the door doN [banging].'

The (marked) triple form is possible, and in that case, the hitting event corresponds to three.

i. Holms-ga doa-o doNdoNdoN(-to) tataita.
Holms-NOM door-ACC IDPH hit
'Holms hit the door doNdoNdoN(-to) [banging].'

[7] The extensibility intensification is also possible in GITAIGO.

these examples. This reading is also derivable even if the ideophone is missing. Another interpretation is also derivable in GISEIGO, whereby a dog's barking approaches the ideal and final state of being waNwaN. Since the ideal state is waNwaN, the barking cannot be described as bow-wow, weMweM, meow, etc. In this reading, the sound of Bella is closer than Alfie to the state of waNwaN in (11a).[8]

(11) a. Bella-ga Alfie yori(mo) (waNwaN(-to)) hoeta.
Bella-NOM Alfie than IDPH barked
'Bella barked more waNwaN than Alfie.'

b. Bella-ga totemo (waNwaN(-to)) hoeta.
Bella-NOM very IDPH barked
'Bella barked very waNwaN.'

To summarize, degree constructions are possible if ideophones can derive a standard of comparison that is abstract and thus not iconic. Since non-phonomime ideophones are based on an abstract notion or are more symbolic, they are available in degree constructions. Phonomime ideophones, by contrast, can build degree constructions as long as a scale is based on some extent that is highly abstract. Their direct or phonomimic interpretations cannot be a standard of comparison because they do not provide scales. I argue that gradability is a diagnostic for the iconicity of ideophones because anti-iconic ideophones are possible in degree constructions. This is reminiscent of the anti-iconicity constraint on Japanese mimetic verb formation by Akita (2009):

(12) The anti-iconicity constraint on Japanese mimetic verb formation: Highly iconic mimetics cannot form verbs (Akita 2009:221).

Adverbial ideophones can be turned into verbs if they are not iconic.[9]

(13) a. dotabata-suru 'make noise, be busy'

[8] If the comparison based on the ideal state is possible, the reading would be a good candidate in analyzing the iconicity in gradability, because the reading would reflect the iconicity in comparisons. This is an interesting topic, but I leave the matter for future research for the reason of space. I thank Kimi Akita and Mark Dingemanse for pointing out this possibility.

[9] About the verb formation by ideophones, see Kiyama & Akita (2015).

b. ♯ doNdoN-suru 'hit' (This is childish and highly colloquial. The sound of hitting is not associated.)

c. * waNwaN-suru

Dotabata has its own abstract meaning, and thus, it can be used as a verb with the light verb -suru, which does not have its own meaning. Since extent reading is not directly derived from doNdoN or waNwaN in comparisons but is derived by combining with their host verbs, verb formation is not possible with these ideophones; they receive interpretations only by combining with their hosts and thus do not provide scales of their own. The availability of degree constructions based on extent and frequency does not indicate anything, but gradability is a good diagnostic of whether some ideophones are iconic or not because the standard they derive implies the existence of an abstract representation that is reflective of symbolic expressions.

3.2 Semantics of Gradable Ideophones

Since some ideophones are gradable, it is possible to hypothesize that they contain degrees (Kennedy 1999; Kennedy & McNally 2005; Kennedy 2007). This subsection provides the semantics of gradable ideophones. For the sake of clarity, I use comparisons where gradable ideophones function as the main predicates.

(14) a. Okuba-ga maeba yori(mo) guragura-suru.
back.tooth-NOM front.tooth than IDPH-do
(Lit.) 'My back tooth seems more guragura [loose] than my front tooth.'

b. Zenkai-no sigoto yori(mo) kutakuta-da.
previous-GEN job than IDPH-be
(Lit.) 'I am more kutakuta [exhausted] than in the last job.'

Assuming that the ideophone guragura contains degrees, the comparative marker yori(mo) denotes degree relations and the verb suru ('do') is semantically vacuous, the meaning of the comparative predicate maeba yori(mo) guragura-suru ('more guragura than my front tooth') in (14a) can be expressed as follows (Kennedy 1999; Kennedy & McNally 2005; Kennedy 2007):

(15) a. $[[\text{guragura}]] = \lambda d\lambda x.\textbf{guragura}(x) = d$

b. $[[d_C \quad \text{yori(mo)}]] = \lambda G\lambda x.\exists d[d \succeq d_C \wedge G(d)(x)]$

c. $[[\text{maeba}]]\ [[\text{yori(mo)}]]\ [[\text{guragura}]]$
$= \lambda x.\exists d[d > \mathbf{d}_{my.front.tooth} \wedge \textbf{guragura}(x) = d]$

In the absence of the comparative phrase hosted by yori(mo), the abstract degree morpheme POS is assumed, which encodes the relation **stnd**, denoting a degree d if it satisfies a standard of comparison for a gradable ideophone **G** with respect to a comparison class determined by contextual information **C**.

(16) a. $[[pos]] = \lambda G\lambda x.\exists d[d \succ \textbf{stnd}(d)(G)(C) \wedge G(d)(x)]$

b. guragura-suru $\Rightarrow$ $[[pos]]([[\textbf{guragura}]])$
$= \lambda x.\exists d[d \succ \textbf{stnd}(d)([[\textbf{guragura}]])(C) \wedge \textbf{guragura}(x) = d]$

Another issue here is what determines the value of the standard for gradable ideophones. Roughly put, the standard for gradable adjectives is either relative (context-sensitive) or absolute (no contextual information). In the latter case, the standard can be either minimal or maximal. A diagnostic for the relative and absolute gradable expressions is to ask what type of degree modifiers are available. Tsujimura (2001) points out that totemo ('very') modifies relative gradable adjectives, which contain contextual information, while sukkari ('entirely') absolute adjectives with a maximal standard. The relevant examples are listed below.

(17) a. totemo {takai, fukai, nagai}
very tall/expensive, deep, long

b. sukkari {kogeta, kootta, kawaita}
entirely burned, frozen, dried

Gradable ideophones, however, are flexible in that they can be modified by either type. Phonomimic ideophones are not compatible with degree modifiers because they do not hold a degree; degree modifiers can combine with verbs via a null degree operator.

(18) a. totemo guragura, sukkari guragura

b. totemo kutakuta, sukkari kutakuta

c. totemo dotabata, sukkari dotabata

d. ♯ totemo doNdoN, ♯sukkari doNdoN

e. ♯ totemo waNwaN, ♯sukkari waNwaN

The meanings of the degree modifiers can be expressed à la Kennedy and McNally (2005); Totemo X is true of an object if the degree to which it is X exceeds a standard even among objects that are X, and sukkari X is true of an object if the degree to which it is maximally X.

(19) a. $[[\text{totemo}]]^C = \lambda G \lambda x.\exists d[d \succ \textbf{stnd}(d)(G)(\lambda y.\ [[pos(G)(y)]]^C) \wedge G(d)(x)]$

b. $[[\text{sukkari}]] = \lambda G \lambda x.\exists d[d = \textbf{max}(S_G) \wedge G(d)(x)]$

The modifiability by the degree modifiers indicates that gradable ideophones are flexible and that their scalar structures are not fixed. Gradable adjectives, by contrast, hold their own scalar structures, which are derived from a conventional standard. Their scales are determined by interpretive economy, which dictates the following: "Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions" (Kennedy 2007). Under the interpretive economy, it follows that ideophones' scales are not conventionalized.

The absolute interpretation is, however, required in their phonologically emphatic forms (e.g., gur̠ragura, kut̠takuta). The emphatic forms resist comparative constructions because a standard must be maximal, and thus, a comparative phrase is redundant, as shown in (20).[10]

[10] 'Expressiveness' might be relevant here. Dingemanse & Akita (2017) claim that as expressiveness increases, grammatical integration decreases. Expressiveness can be measured by phonetic emphasis or frequent gestures, and the degree of grammatical integration is determined by whether some items can be used as core items–subjects, objects, predicates, etc. A relevant example is shown below:

i. Sonoɯtʃikawara-ga gatʃagatʃag-gatʃa:t-to otʃi-te kɯ-ɾɯ.
soon.tile-NOM IDPH.PM1.VL.SR.VOICELESS-QUOT fall-CONJ come-NPST
'Then, the roofing tiles drop down on us with a loud clattering noise [both hands loosely open, palms down, moving slightly up and down in front of the speaker's chest, synchronized with the production of the ideophone].'

ii. Mo: bo:hate: giɾigiɾi-des-ɯ.
already breakwater IDPH-COP.POL-NPST
'[The sea level] was already almost reaching the breakwater.' (Dingemanse & Akita 2017)

(20) a. ?? Okuba-ga maeba yori(mo) gu**r̲ragura**-suru.

b. ?? Zenkai-no sigoto yori(mo) ku**t̲**takuta-da.

To summarize, some ideophones are gradable and participate in grammatical integration (i.e., build comparative constructions). Scalar structures for ideophones are flexible in that their standard can be either relative or absolute. The maximum standard is required for emphatic ideophones because the extremeness is expressed by the ultimate or maximum standard. Since the standard for emphatic ideophones is already maximal, the standard marker hosted by yori(mo) leads to redundancy, resisting comparative constructions.

4 Conclusion

I have shown that there is a gradience among ideophones. Some ideophones are gradable and symbolic. These ideophones are not special semantically but are peculiar from a morphological or phonological point of view. Other ideophones are iconic and have strong depictive flavor, resisting grammatical integration because they are unavailable as main predicates or do not provide a standard of comparison. Since gradability is based on scales, this will shed light on how special, or not, some ideophones are.

References

AKITA, KIMI, 2009. *A Grammar of Sound-Symbolic Words in Japanese*. Kobe University dissertation.

——, & TAKESHI USUKI. 2016. A constructional account of the 'optional' quotative marking on Japanese mimetics. *Journal of Linguistics* 52.245–275.

BOLINGER, DWIGHT. 1968. *Aspects of Language*. New York: Harcourt, Brace and World.

Here, the ideophone gatʃagatʃag-gatʃa:t-to is high in expressiveness because it contains phonetic emphasis and is accompanied by a gesture, while it is low in grammatical integration and thus is a modifier. The ideophone giɾigiɾi-des-ɯ, by contrast, is low in expressiveness but high in grammatical integration. Hence, it can serve as a main predicate.

Since the emphatic ideophones have a high degree of expressiveness, they resist combining with comparative phrases, which constitute constructions with high grammatical integration.

BURCH, ROBERT. 2017. Charles sanders peirce. In *The Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta. Metaphysics Research Lab: Stanford University.

DIFFLOTH, GÉRARD. 1972. Notes on expressive meaning. In *Chicago Linguistic Society*, volume 8, 440–447.

DINGEMANSE, MARK, 2011. *The Meaning and Use of Ideophones in Siwu*. Max Plank Institute for Psycholinguistics dissertation.

——. 2012. Advnaces in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6.654–672.

——. 2015. Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language* 39.946–970.

——. 2017. Expressiveness and system integration: On the typology of ideophones, with special reference to Siwu. *STUF – Language Typology and Universals* 70.363–384.

——. 2018. Redrawing the margins of language: Lessons from research on ideophones. *Glossa* 3.1–30.

——, & KIMI AKITA. 2017. An inverse relation between expressiveness and grammatical integration: On the morphosyntactic typology of ideophones, with spetial reference to Japanese. *Journal of Linguistics* 53.501–532.

DOKE, CLEMENT MARTYN. 1935. *Bantu Linguistic Terminology*. London: Longmans, Green and Co.

GÜLDEMANN, TOM. 2008. *Quotative Indexes in African Languages: A Synchronic and Diachronic Survery*. Berlin: Mouton de Gruyter.

KAKEHI, HISAO, & IKUHIRO TAMORI. 1993. *Onomatopoeia-Gion/Gitaigo-no Rakuen*. Keisoshobo.

KENNEDY, CHRISTOPHER. 1999. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. New York: Garland Publishers.

——. 2007. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30.1–45.

——, & LOUISE MCNALLY. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81.345–381.

KINDAICHI, HARUHIKO. 1978. *Giongo, Gitaigo Ziten (Mimetics dicitonary)*. Tokyo: Kadokawa.

KITA, SOTARO. 1997. Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35.379–415.

KIYAMA, NAOKI, & KIMI AKITA. 2015. Gradability and mimetic verbs in Japanese: A frame-semantic account. In *Berkeley Linguistics Society*, ed. by Anna Jurgensen, Hannah Sande, Spencer Lamoureux, Kenny Baclawski, & Alison Zerbe, volume 41, 245–266.

MOORE, COLETTE. 2015. An ideological history of the Englsih term onomatopoeia. In *Studies in the History of the Englsih Language*, ed. by Michael Adams, Laurel J. Brinton, & R.D. Fulk, 307–320. Walter de Gruyter.

SAWADA, OSAMU. 2013. The comparative morpheme in Modern Japanese: Looking at the core from 'outside'. *Journal of East Asian Linguistics* 19.191–226.

TAMORI, IKUHIRO, & LAWRENCE SCHOURUP. 1999. *Onomatopoe: Keitai-to imi (Onomatopoeia: Form and Meaning)*. Tokyo: Kurosio Publishers.

TSUJIMURA, NATSUKO. 2001. Degree words and scalar strucutre in Japanese. *Lingua* 111.29–52.

WESTERMANN, DIEDRICH HERMANN. 1907. *Grammatik der Ewe-Sprache*. Berlin: Dietrich Reimer.

Suppletive allomorphy conditioned by humbleness in Korean

SOO-HWAN LEE & MINJUNG KIM
Sogang University[*]

1 Introduction

This paper demonstrates that a postsyntactic derivational mismatch exists between the two phonological elements, *ney* 'yes.HUM' and *ce* 'I.HUM', which are both conditioned by the humble feature, [+hum], in Korean. With the goal of verifying such mismatch, this paper sheds light upon the role of fusion which is known to be a postsyntactic apparatus utilized within the framework of Distributed Morphology (Halle & Marantz 1993, 1994). In addition, this work claims that the (in)applicability of fusion ultimately determines the (un)availability of multiple humble feature linearization, $[+\text{hum}]_1 \frown [+\text{hum}]_2$, in contexts where the conventional implicature (CI) of politeness is at play (see Potts 2005 on CI *expressives*). While *ney* and *ce* share the commonality of conveying politeness within a given discourse, the ways in which they are structurally organized seem to differ as the attachment of an additional humble feature, $[+\text{hum}]_2$, distinguishes the former from the latter in terms of grammatical well-formedness. In order to put forward a satisfactory explanation for such mismatch, this paper scrutinizes the derivations of *ney* 'yes.HUM' and *ung* 'yes' as well as *ce* 'I.HUM' and *na* 'I' which are relevant to suppletive allomorphy.

2 Humbleness in Korean

Prior to analyzing suppletion conditioned by humbleness in Korean, let us observe how the humble feature, [+hum], may be realized within a given sentence structure. It is well known in literature that Korean makes use of the honorific morpheme, *si*. Similar to the ways in which the honorific feature, [+hon], is realized as the phonological form, *si*, the humble feature, [+hum], is

---

[*]This work was supported by CORE(Initiative for College of Humanities' Research and Education) and funded by the Ministry of Education and the National Research Foundation of Korea. We would like to thank the audience at BLS 44 for their valuable comments. We would also like to express our thanks to Inkie Chung for his endless support. All errors are our own.

also capable of being realized as the phonological entity, *yo*. Here, the distinction between the honorific feature, [+hon], and the humble feature, [+hum], comes to light when the following examples are given consideration:[1]

(1)	a.	ese	o(wa)-**si(e)-yo**
		quickly	come-**HON-HUM**
		'Welcome.' (very polite)	
	b.	ese	o(wa)-**yo**
		quickly	come-**HUM**
		'Welcome.' (polite)	
	c.	ese	o(wa)
		quickly	come
		'Welcome.'	

As shown in 1a and 1b, the phonological form, *yo*, bearing the feature, [+hum], may surface with or without the presence of the honorific morpheme, *si* [+hon]. In cases where [+hum] is realized without the honorific morpheme as shown in 1b, the degree of politeness stays low when compared to that of 1a hosting both the honorific feature, *si* [+hon], as well as the humble feature, *yo* [+hum]. Nevertheless, it is not difficult to understand that the realization of *yo* [+hum] alone is enough to trigger politeness. Such notion can be verified when 1b is compared to 1c which bears no phonological element indicating politeness. Considering such dissimilarities, we discern that the attachment of a humble feature expresses a form of politeness while the attachment of a humble feature in addition to an honorific feature gives rise to an elevated measure of politeness. Thus, it can be said that the humble feature, [+hum], holds onto the effect of either introducing or magnifying politeness in Korean. In further pursuit of analyzing the humble feature, [+hum], this paper attempts to clarify the element's apparent involvement with suppletive allomorphy in the following section.

[1] In this paper, we use of the following the abbreviations: ACC = accusative; D = default; DECL = declarative; DM = distributive marker; HON = honorific; HUM = humble; LOC = locative; NEG = negation; NOM = nominative; PST = past.

3 Answering polar questions in Korean

In light of analyzing the ways in which the feature, [+hum], interact with suppletive allomorphy, we delve into the morphology of polar answers, namely *ani* 'no' and *ung* 'yes', in Korean. Consider the following data which illustrate a polar question followed by four possible ways of answering the question:

(2)	Cinha-ka	ppang-ul	mek-ess-ni?
	Cinha-NOM	ppang-ACC	eat-PST-Q
	'Did Cinha eat the bread?'		

(3)	a.	ani	b.	ani-yo	c.	ung	d.	ney
		No		no-HUM		yes		yes.HUM
		'No.'		'No.' (polite)		'Yes.'		'Yes.' (polite)

3.1 Negative responses

In terms of providing a negative response to the question presented in 2, there are two basic ways of doing so which are briefly shown in 3a and 3b. The difference between the two crucially depends on whether politeness is involved or not between the speaker and the addressee who participate in a given conversation. Well-formed negative responses realized within full sentence structures appear as the following data:

(4)	a.	**ani**	Cinha-ka	ppang-ul	an	mek-ess-e
		no	Cinha-NOM	ppang-ACC	NEG	eat-PST-D
		'No, Cinha did not eat the bread.'				
	b.	**ani-yo**	Cinha-ka	ppang-ul	an	mek-ess-e-**yo**
		no-HUM	Cinha-NOM	ppang-ACC	NEG	eat-PST-D-**HUM**
		'No, Cinha did not eat the bread.' (polite)				

As opposed to 4a, 4b hosts the humble feature, *yo*, in a sentence-final position in order to signify a sense of civility. This is in line with what we have observed in 1b. Interestingly, however, an additional appearance of the morpheme, *yo*, can be observed in 4b. The multiple appearance of *yo* which hosts the morphosyntactic feature [+hum] is explained through what is referred to as the sentence-medial morpheme, *yo* (Yim 2012).

(5)	ani-**yo**	Cinha-ka-**yo**	ppang-ul-**yo**	an	mek-ess-e-**yo**
	no-**HUM**	Cinha-NOM-**HUM**	ppang-ACC-**HUM**	NEG	eat-PST-D-**HUM**

‘No, Cinha did not eat the bread.’ (polite)

As long as the sentence-final morpheme, *yo*, is properly assigned, the insertion of the sentence-medial *yo* is possible as shown in 5. Adopting the notion that multiple existence of humble features in Korean is possible, we realize that *ani-yo* does not crash in derivation. Further evidence on the conditioning of the sentence-medial element, *yo*, can be seen in the following data:

(6)	*ani-**yo**	Cinha-ka	ppang-ul	an	mek-ess-e
	no-**HUM**	Cinha-NOM	ppang-ACC	NEG	eat-PST-D

‘No, Cinha did not ate the bread.’ (polite)

The realization of *ani-yo* in 6 is not possible because the sentence final *yo* does not surface. Such notion stays in line with Yim’s (2012) argument that the sentence-final element must be present when their sentence-medial counterparts are in place. Putting together the evidence collected so far, we realize that the phonological entity, *ani-yo* ‘no-HUM’, consists of two abstract features, namely √NO and [+hum], prior to Vocabulary Insertion.

(7)	√NO	↔	*ani*
	[+hum]	↔	*yo* / *ani* _______

During Vocabulary Insertion, the phonological realizations of *ani* ‘√NO’ and *yo* ‘[+hum]’ take place consecutively as shown in 7. Here, the humble feature, [+hum], does not trigger suppletive allomorphy on √NO as there is no idiosyncrasy in form. However, it will be proven otherwise for the items conveying positive responses in the following subsection.

3.2 Positive responses

In addition to the overt morphological form, *yo*, a humble feature gives rise to suppletive allomorphy in Korean. Examples of suppletive allomorphy can be observed through the systematic alternation between the phonological components, *ney* 'yes.HUM' and *ung* 'yes'. When a polar question such as 2 is provided, there are two grammatical ways of forming a positive response. The first option is to use *ney* in order to convey politeness, whereas the second option is to use *ung* without conveying such implicit connotation.

(8)	a.	ung	Cinha-ka	ppang-ul	mek-ess-e
		yes	Cinha-NOM	ppang-ACC	eat-PST-D
		'Yes, Cinha ate the bread.'			
	b.	**ney (*ung-yo)**	Cinha-ka	ppang-ul	mek-ess-e-**yo**
		yes.HUM	Cinha-NOM	ppang-ACC	eat-PST-D-**HUM**
		'Yes, Cinha ate the bread.' (polite)			

Surprisingly enough, we see that the attachment of the phonological element, *yo*, to *ung* 'yes' is not a well-formed realization as pointed out in 8b. This conflicts with the negative response shown in 4b in terms of how the feature, [+hum], may be realized. Clearly, there arises a mismatch between *ani-yo* 'no-HUM' and *ney*(**ung-yo*) 'yes-HUM' in terms of grammatical well-formedness. As a way of abandoning the usage of **ung-yo* 'yes-HUM' completely, the realization of *ney* 'yes.HUM' is utilized which runs counter to the idea of morpheme agglutination exemplified by *ani-yo* 'no-HUM' in 4b. Despite the fact that *ney* 'yes.HUM' and *ani-yo* 'no-HUM' differ in terms of hosting the overt element, *yo*, both entities convey the same type of conventional implicature relevant to politeness. Such idea is adopted from Huang (2014) as well as Potts (2005, 2007). (The figure, +>, indicates the reading of a conventional implicature).

(9)	a.	ani-yo
		no-HUM

'No.' (polite)
+> The speaker is socially inferior to or distant from the addressee

b. ney
yes.HUM
'Yes.' (polite)
+> The speaker is socially inferior to or distant from the addressee

Similar to the negative response shown in 9a, the positive response in 9b clearly conveys a sense of politeness. Taking this into account, we realize that the only noticeable difference of the two is the mismatch between their dissimilar morphological forms. In this regard, *ney* 'yes.HUM' is morphologically idiosyncratic, because it does not stay in line with the prediction that the morpheme, *yo*, should appear. Nevertheless, the fact that *ney* 'yes.HUM' conveys a conventional implicature identical to that of *ani-yo* 'no-HUM' must be maintained. Otherwise, there is no way of accounting for the existence of a polite reading available in both 9a and 9b. So as to resolve this apparent dilemma, this paper argues that *ney* 'yes.HUM' is a case of suppletive allomorphy conditioned by the morphosyntactic feature, [+hum], which does not surface as the phonological form, *yo*. Following this string of logic, we postulate that the catalyst of suppletion is [+hum] which triggers allomorphy on the target, √YES. Subsequently, the catalyst itself receives a null exponent within the environment where the allomorph, *ney*, is phonologically realized. The gist of the proposal is illustrated below:

(10)	√YES	↔	*ney*	/	_______ [+hum]
		↔	*ung*	/	elsewhere
	[+hum]	↔	Ø	/	*ney* _______
		↔	*yo*	/	elsewhere

As demonstrated in 10, the derivations of the forms *ney* 'yes.HUM' and *ung* 'yes' are made possible during the stage of Vocabulary Insertion. Further note that the phonologically null element, Ø, which corresponds to [+hum], is assigned only after the form, *ney*, because such assertion is able to also account for the realization of *ani-yo* 'no-HUM' where the phonologically overt element, *yo*, follows the phonological entity, *ani*. Hence, consistency in terms of morphological ordering is maintained in both cases. Additional examples of suppletive allomorphy will be introduced in the following section which focuses on the realization of first person pronouns in Korean.

4 First person pronouns

It is well known in literature that first person pronouns in Korean are morphologically sensitive to the discourse influence of politeness. As for first person singular pronouns, the alternation between *na* 'I' and *ce* 'I.HUM' takes place, whereas for first person plural pronouns, the alternation between *wuli* 'we' and *cehuy* 'we.HUM' takes effect. With the goal of explaining such alternations, this paper further investigates the role of the humble feature, [+hum], on Korean suppletive allomorphy.

4.1 Expressing first person singular pronouns

The following data provides an outline of how suppletive allomorphy of the 1st person singular pronoun in Korean is conditioned by [+hum]. Bear in mind that *nay* 'I' and *cey* 'I.HUM' are just morphophonological alternants of *na* 'I' and *ce* 'I.HUM'.

(11) a. **nay**-ka Inwu-lul cap-ass-e
I-NOM Inwu-ACC catch-PST-D
'I caught Inwu.'

b. **cey**-ka Inwu-lul cap-ass-e-**yo**
I.HUM-NOM Inwu-ACC catch-PST-D-**HUM**
'I caught Inwu.' (polite)

As shown in 11, the overt exponent, *ce*, is realized when the sentence-final humble feature is properly assigned, whereas the exponent, *na*, is realized when there is no indication of the humble feature. Continuing on with the discussion on suppletive allomorphy, we postulate that the appearance of *na* or *ce* is conditioned by the presence or the absence of [+hum]. Consider the following data which is in line with this analysis:

(12) a. ***cey-ka** Inwu-lul cap-ass-e
I.HUM-NOM Inwu-ACC catch-PST-D
'I caught Inwu.' (polite)

b. ***ney** Chelswu-ka Inwu-lul cap-ass-e
yes.HUM Chelswu -NOM Inwu-ACC catch-PST-D
'Yes, Chelswu caught Inwu.' (polite)

Note that *ce* cannot be realized without the sentence-final element, *yo*, which patterns with the realization of the allomorph, *ney* 'yes.HUM', discussed in section 3. Not only does this speak in favor of Yim's (2012) argument regarding the distribution of *yo*, but it also suggests that the interchangeable use of *ce* 'I.HUM' and *na* 'I' is allomorphic which is identical to the pattern displayed between *ney* 'yes.HUM' and *ung* 'yes' in (8). Moreover, the shift between *na* and *ce* is suppletive due to the notion that the two exponents are phonologically irrelevant to one another. Thus, we are able to capture the systematic alternation between the two forms through a design running on the basis of competition introduced in Distributed Morphology:

(13) [+1st, +sg] ↔ *ce* / _______ [+hum]
↔ *na* / elsewhere

In accordance with such view, 13 summarizes the rule accurately predicting the phonological realizations of the first person singular pronoun, *na* and *ce*, depending on whether [+hum] participates in suppletive allomorphy or not. Recall that the suppletive target precedes the trigger, [+hum], for reasons which have been discussed in section 3 (e.g, *ani-yo* 'no-HUM').

4.2 Expressing first person plural pronouns

The line of argument provided for the alternation between *na* and *ce* applies to first person plural pronouns in Korean as well. In other words, making a systematic prediction for the alternation between the plural pronoun forms, *cehuy* 'we.HUM' and *wuli* 'we', is possible depending on the presence or the absence of [+hum].

(14)	a.	**wuli**-ka	Inwu-lul	cap-ass-e
		we-NOM	Inwu-ACC	catch-PST-D
		'We caught Inwu.'		
	b.	**cehuy**-ka	Inwu-lul	cap-ass-e-**yo**
		we.HUM-NOM	Inwu-ACC	catch-PST-D-**HUM**
		'We caught Inwu.' (polite)		

As shown in 14a, the form, *wuli*, is realized when the humble feature is absent. On the contrary, the form, *cehuy*, is realized when the sentence-final element hosting [+hum] is present as shown in 14b. Surely, this is in harmony with what we have seen for the phonological alternation between *na* 'I' and *ce* 'I.HUM' in 11.

(15)	a.	**wuli**	hayngpokhan	enehakja-tul
		we	happy	linguist-DM
		'We happy linguists'		
	b.	**cehuy**	hayngpokhan	enehakja-tul
		we.HUM	happy	linguist-DM
		'We happy linguists' (polite)		

Further implication arises when we look into pronoun-noun constructions (PNCs) in Korean (see Choi 2014 for more detail). As it is demonstrated in 15, the alternation between *wuli* and *cehuy* can be observed even within Korean PNCs. As expected, the nominal structure in 15b is in need of a sentence-final humble feature so that the allomorph, *cehuy*, may surface during Vocabulary Insertion. Moreover, within the environment where [+hum] is absent, only the phonological realization of *wuli* 'we' is possible as in 15a. Following this string of logic, it becomes once again valid to assert that the systematic distribution of *wuli* and *cehuy* runs on the basis of competition.

(16)	[+1st, -sg]	↔	*cehuy*	/ _______ [+hum]
		↔	*wuli*	/ elsewhere

To briefly summarize this part of the section, suppletive allomorphy of the first person plural pronouns in Korean is conditioned by the humble feature, [+hum]. Such analysis patterns with the ways in which the positive response, *ney* ‘yes.HUM’ as well as the first person singular pronoun, *ce* ‘I.HUM’, are conditioned. As it is illustrated in 16, the two competing elements, *wuli* and *cehuy*, may be assigned to their proper environments via the mechanism of Vocabulary Insertion.

5 The (in)applicability of fusion

In this section, we scrutinize the morphosyntactic mismatch displayed during the derivations of the suppletive forms, *ney* ‘yes.HUM’ and *ce*/*cehuy* ‘I.HUM/we.HUM’. Here, the first person pronouns, *ce* and *cehuy*, are grouped together into a single category since their morphosyntactic characteristics are not different. So as to verify the mismatch between *ney* and *ce*/*cehuy*, this paper introduces the postsyntactic operation known as fusion within the literature of Distributed Morphology. The definition of the operation is provided below:

(17) Fusion takes two terminal nodes that are sisters under a single category node and fuses them into a single terminal node. (Halle & Marantz 1993)

Depending on whether the postsyntactic apparatus, fusion, is applied or not, the (un)availability of the multiple consecutive linearization of $[+\text{hum}]_1 \frown [+\text{hum}]_2$ can be anticipated. Here, it is worth mentioning that the multiple consecutive realization of the humble features, $[+\text{hum}]_1 \frown [+\text{hum}]_2$, is not possible when the features are stacked up compositionally which enables the secondary feature, $[+\text{hum}]_2$, to gain access to the primary feature, $[+\text{hum}]_1$.

(18) a. *Cinha-ka-**yo-yo** ppang-ul mek-ess-e-yo
Cinha-NOM-**HUM-HUM** bread-ACC eat-PST-D-HUM
‘Cinha ate the bread.’ (polite)

b. *Cinha-ka-yo ppang-ul mek-ess-e-**yo-yo**
Cinha-NOM-HUM bread-ACC eat-PST-D-**HUM-HUM**
‘Cinha ate the bread.’ (polite)

As illustrated in 18a, the consecutive linearization of *yo-yo* ‘HUM-HUM’ via the distribution of the sentence-medial elements, $[+\text{hum}]_1$ and $[+\text{hum}]_2$, crashes in derivation, because the primary humble feature is visible to the secondary feature. The same argument can be made for 18b where the consecutive linearization of *yo-yo* 'HUM-HUM' in the sentence-final position is unavailable due to the visibility of $[+\text{hum}]_1$ to $[+\text{hum}]_2$.

(19)	***ney-yo**	Cinha-ka	cip-ey	ka-yo
	yes.**HUM-HUM**	Cinha-NOM	house-LOC	go-HUM
	'Yes, Cinha is going home.' (polite)			

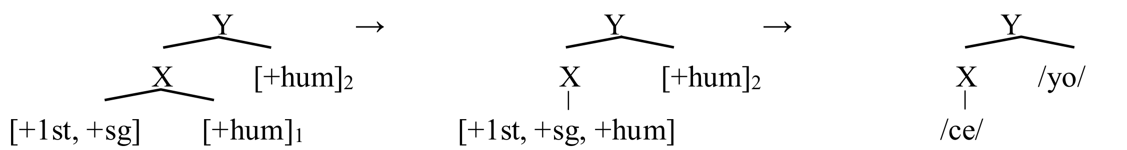
Such line of reasoning for the invalidity of $[+\text{hum}]_1⌒[+\text{hum}]_2$ gains further significance when the ill-formedness of **ney-yo* 'yes.HUM-HUM' is given consideration in 19. Here, $[+\text{hum}]_2$ once again gains access to $[+\text{hum}]_1$ during the derivation of suppletive allomorphy. Hence, the ungrammaticality of **ney-yo* 'yes.HUM-HUM' follows easily. Quite interestingly, however, not all incidents of suppletive allomorphy conditioned by humbleness override the realization of $[+\text{hum}]_1⌒[+\text{hum}]_2$. In fact, the pronominal allomorphs, *ce* 'I.HUM' and *cehuy* 'we.HUM', discussed in section 4 are compatible with the secondary humble feature, $[+\text{hum}]_2$, which is later realized as *yo*.

(20)	a.	**ce-yo**	cip-ey	ka-yo
		I.**HUM-HUM**	house-LOC	go-HUM
		'I am going home.' (polite)		
	b.	**cehuy-yo**	cip-ey	ka-yo
		we.**HUM-HUM**	house-LOC	go-HUM
		'We are going home.' (polite)		

As demonstrated in 13 and 16, the suppletive forms, *ce* 'I.HUM' and *cehuy* 'we.HUM', appear only when the target, [+1st, ±sg], immediately precedes the trigger, [+hum]. Unsurprisingly, the suppletive forms, *ce* and *cehuy*, also convey the same type of conventional implicature due to the very presence of [+hum]. At this point, it is more than crucial to understand that the sentence-medial element, *yo*, hosting $[+\text{hum}]_2$, is able to reside next to *ce* 'I.HUM' and *cehuy* we.HUM' which evidently gives rise to the multiple consecutive linearization of $[+\text{hum}]_1⌒[+\text{hum}]_2$

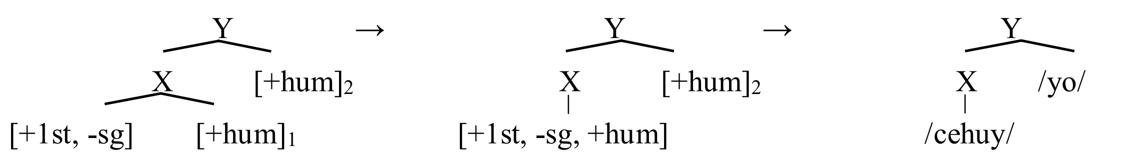
illustrated in 20. In order to put forward a valid explanation for the multiple consecutive linearization of $[+\text{hum}]_1 \frown [+\text{hum}]_2$, there must be a way of making $[+\text{hum}]_1$ *invisible* to $[+\text{hum}]_2$ so that the realization of $[+\text{hum}]_1 \frown [+\text{hum}]_2$ is possible. In fact, Nunes (2004) proposes a similar analysis on utilizing the postsyntactic operation, fusion, in order to disenable PF-deletion from taking effect on certain intermediate *wh*-copies during successive-cyclic *wh*-movements. Here, we argue that certain suppletive allomorphs in Korean also undergo fusion in order to make the multiple consecutive linearization of $[+\text{hum}]_1 \frown [+\text{hum}]_2$ possible within their given morphosyntactic structures. Consider the following procedures of derivation for *ce* and *cehuy*:

(21) a. (Post)syntax Fusion Vocabulary Insertion



b. [+1st, +sg, +hum] ↔ *ce*
[+1st, +sg] ↔ *na*
[+hum] ↔ *yo*

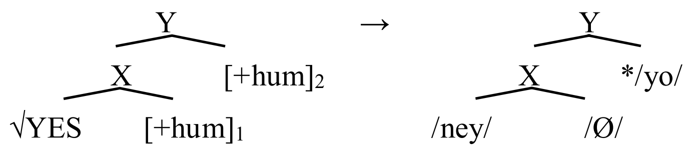
(22) a. (Post)syntax Fusion Vocabulary Insertion



b. [+1st, -sg, +hum] ↔ *cehuy*
[+1st, -sg] ↔ *wuli*
[+hum] ↔ *yo*

By means of presenting a theoretical account on the well-formedness of *ce-yo* 'I.HUM-HUM' and *cehuy-yo* 'we.HUM-HUM', the postsyntactic apparatus, fusion, is employed in both 21 and 22. As a result, $[+\text{hum}]_1$ is left undetected by $[+\text{hum}]_2$ which eventually enables $[+\text{hum}]_1 \frown [+\text{hum}]_2$ to surface without any given restrictions being violated. As an additional consequence for following such analysis, we apprehend that fusion does not take effect in situations where $[+\text{hum}]_1 \frown [+\text{hum}]_2$ devastates a given derivation. For this reason, this paper argues that not all derivational processes of suppletive allomorphy conditioned by humbleness undergo fusion. Consider the following procedural derivation of *ney* 'yes.HUM'.

(23) a. (Post)syntax → Vocabulary Insertion



b.

√YES	↔	*ney*	/	_______ [+hum]
	↔	*ung*	/	elsewhere
[+hum]	↔	Ø	/	*ney* _______
	↔	*yo*	/	elsewhere

As for the morphological form, **ney-yo* 'yes.HUM-HUM', in 19 and 23, the application of fusion must be ruled out, for the otherwise scenario would incorrectly predict the invisibility of $[+\text{hum}]_1$ to its hierarchically dominant feature, $[+\text{hum}]_2$. Following this analysis, the application of the postsyntactic device, fusion, would overestimate the grammaticality of **ney-yo*. This, in turn, would further suggest that fusion is not a guaranteed option even if the trigger, [+hum], and target, √YES, are in a sister relation. In these regards, fusion should not be involved in the derivation of *ney*, which consequently leads to an apparent mismatch in derivation between *ce*/*cehuy* in 21 and 22 and *ney* in 23.

6 Conclusion

To sum up the discussion made thus far, a derivational mismatch between the suppletive forms, *ce*/*cehuy* 'I.HUM/we.HUM' and *ney* 'yes.HUM', has been proposed on the basis of whether the postsyntactic operation known as fusion takes effect prior to Vocabulary Insertion at PF. To be more specific, the former is compatible with fusion, whereas the latter is not. In further pursuit of verifying this proposal, it has been suggested that the (in)applicability of fusion determines the (un)availability of the multiple consecutive linearization of [+hum]$_1$⁀[+hum]$_2$, which explains the morphosyntactic discrepancy between *ce-yo* 'I.HUM-HUM'/*cehuy-yo* 'we.HUM-HUM' and **ney-yo* 'yes.HUM-HUM'. All in all, this work has proven that not all cases of suppletive allomorphy conditioned by the humble feature, [+hum], in Korean are subject to a single derivational process especially when the (in)applicability of the postsyntactic operation, fusion, is given consideration.

REFERENCES

BOBALJIK, JONATHAN and HEIDI HARLEY. 2017. Suppletion is local: Evidence from Hiaki. *The Structure of Words at the Interfaces*, edited by H. Newell, M. Noonan, G. Piggot & L. Travis, 141-159. Oxford: OUP.

CHOI, JAEHOON. 2014. *Pronoun-noun constructions and the syntax of DP*. Ph.D. Dissertation, University of Arizona dissertation.

CHOI, JAEHOON and HEIDI HARLEY. 2017. Node sprouting and root suppletion in Korean verbal morphology. *Talk at GLOW in Asia XI*.

CHOMSKY, NOAM. 1995. *The Minimalist Program*. Cambridge, Mass.: The MIT Press.

CHUNG, INKIE. 2007. Suppletive Negation in Korean and Distributed Morphology. *Lingua* 117, 95-148.

CHUNG, INKIE. 2009. Suppletive Verbal Morphology in Korean and the Mechanism of Vocabulary Insertion. *Journal of Linguistics* 45, 533-567.

DEAL, AMY ROSE. 2016. Plural exponence in the Nez Perce DP: A DM analysis. *Morphology*, 26(3-4), 313-339.

EMBICK, DAVID. 2010. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT.

EMBICK, DAVID. 2015. *The Morpheme: A Theoretical Introduction* (Vol. 31). Boston & Berlin: Mouton de Gruyter.

HAEGEMAN, LILIANE and VIRGINIA HILL. 2013. The syntacticization of discourse. *Syntax and Its Limits*, 370-390.

HALLE, MORRIS and ALEC MARANTZ. 1993. Distributed Morphology and the Pieces of Inflection. In Ken Hale and Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 111-176. Cambridge, MA: MIT Press.

HALLE, MORRIS and ALEC MARANTZ. 1994. Some Key Features of Distributed Morphology. *MIT Working Papers in Linguistics* 21, 275-288.

HANKAMER, JORGE and LINE MIKKELSEN. 2018. Structure, architecture, and blocking. *Linguistic Inquiry* 49(1): 61-84.

HUANG, YAN. 2014. *Pragmatics*. Oxford: Oxford University Press.

KIM, JAIEUN, and INKIE CHUNG. 2015. Hankwuke Jontay Cheykyeyey Tayhan Thonghapcek Pwunsanhyengthaylon Pwunsek (=A Unified Distributed Morphology Analysis of Korean Honorification Morphology). *Studies in Generative Grammar*, 25, 631-650.

KIM, MINJUNG and SOO-HWAN LEE. 2017. Passive suppletion of mac- 'to be hit' in Korean. In the Proceedings of *the 43rd Annual Meeting of Berkeley Linguistics Society* 43(2), 5-10.

MIYAGAWA, SHIGERU. 2012. Agreements that occur mainly in the main clause, *Main Clause Phenomena: New Horizons*, edited by L. Aelbrecht, L. Haegeman & R. Nye, 79-111. Amsterdam/Philadelphia: John Benjamins.

NUNES, JAIRO. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

POTTS, CHRISTOPHER. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.

SPEAS, PEGGY and CAROL TENNY. 2003. Configurational Properties of Point of View Roles. *Linguistics Today Asymmetry in Grammar*, 315-344.

TROMMER, JOCHEN. 1999. Morphology consuming syntax' resources: Generation and parsing in a minimalist version of Distributed Morphology. Ms., Universität Potsdam, Potsdam. In *Proceedings or the ESSLI Workshop on Resource Logics and Minimalist Grammars*. Utrecht, August 1999.

YIM, CHANGGUK. 2012. Fragment answers containing -yo in Korean: new evidence for the PF deletion theory of ellipsis. *Linguistic Inquiry* 43(3): 514-518.

A feature-based analysis of the Ch'ol (Mayan) person paradigm

CAROL-ROSE LITTLE
*Cornell University**

1 Introduction

This paper investigates first person plurals in Ch'ol, a Mayan language of southern Mexico. The Ch'ol absolutive markers are given in Table 1.[1]

Table 1: Absolutive markers in Ch'ol

1	2	3	1 INCL	1 EXCL	2 PL	3 PL
-oñ	*-ety*	Ø	***-oñ*** $\boxed{=la}$	***-oñ*** $\boxed{=loj}$ ***-oñ***	*-ety=* $\boxed{la}$	*-ob*

At first glance it seems that the clitic *=la* (in boxes) shows up in the inclusive plural and second person plural forms. However, a closer look at the paradigm reveals that the exclusive plural form *oñ=loj-oñ* is derived by adding *-oñ*, the first person form, to the inclusive form *oñ=la*. There is vowel assimilation of *a* to *o* in the participant pluralizer *la*, and a [h] (orthographically *j*) is added to break up the vowel hiatus in the exclusive form. Effectively, the exclusive plural is derived by adding the first person marker to the inclusive form.

This data brings up two questions: (i) Why does the exclusive form contain the inclusive form? (ii) Why is it that Ch'ol derives the exclusive by repeating the first person marker *oñ* to the inclusive form? In this paper, I aim to answer these questions. To do so, I

*I thank the patience of the Ch'ol speakers I have worked with: namely the Arcos López family in San Miguel and Morelia Vázquez Martínez and Virginia Martínez Vázquez in El Campanario. I thank Miloje Despić, Jessica Coon, Sarah Murray, Mary Moroney, the Cornell Semantics Group and the audience at BLS 44 for comments and discussion. Unless otherwise marked, data comes from my fieldwork in Chiapas, Mexico. Any errors are my own.

[1]Glosses: 1 = first person; 2 = second person; 3 = third person; A = Set A markers (ergative/possessive); B = Set B markers (absolutive); IMPF = imperfective; IV = intransitive verb; PART = participant; PL = plural; PREP = preposition; PROG = progressive; PRF = perfective aspect; REA = realis. Ch'ol uses a Spanish-based orthography: ' = [ʔ]; ä = [ɨ]; b = [ɓ] ch = [tʃ]; j = [h]; ñ = [ɲ]; ty = [tʲ]; x = [ʃ]; y = [j]; C' = ejective consonant.

take a closer look into the morphology of the two first person plural forms and the contexts in which these two forms occur (§2). Based on empirical generalizations from the usage of the two first person plural forms, I recategorize the inclusive as a general first person plural and the exclusive as a specified exclusive form that explicitly excludes the hearer. As the data will provide evidence for below, the exclusive form is both morphologically and semantically more complex. In §3, I formalize the person paradigm in Ch'ol with binary features (e.g. Noyer 1992, Bobaljik 2008, Watanabe 2013). I argue that a binary feature approach with a [±hearer] feature can better capture the Ch'ol data as it can explicitly exclude the hearer from the representation. In §4, I argue that the Ch'ol data poses challenges for privative feature approaches, like feature geometries (e.g. Harley & Ritter 2002, Cowper & Hall 2005) as these privative systems do not have a way to explicitly exclude a [hearer] feature. I summarize and conclude in §5.

2 Morphology and distribution of first person forms

2.1 Morphology of person markers in Ch'ol

Person in Ch'ol is indexed on the predicate (noun or verb), with two sets of person markers known as set A for ergative/possessive forms and set B for absolutive forms. Set A and set B morphemes are given in Table 2.

Table 2: Person markers and pronouns in Ch'ol

	1	2	3	1 INCL	1 EXCL	2 PL	3 PL
Set A	*k-/j-*	*a(w)-*	Ø	*k-/j-*... $\boxed{=la}$	*k-/j-*... $\boxed{=loj}$***-oñ***	*a(w)-*... $\boxed{=la}$	*-ob*
Set B	***-oñ***	*-ety*	Ø	***-oñ***$\boxed{=la}$	***-oñ***$\boxed{=loj}$***-oñ***	*-ety=*$\boxed{la}$	*-ob*
Pron.	*joñ-**oñ***	*jaty-ety*	*jiñ*	*joñ-**oñ***$\boxed{=la}$	*joñ-**oñ***$\boxed{=loj}$***-oñ***	*jaty-ety=*$\boxed{la}$	*jiñ-ob*

Human-denoting nouns and some animate nouns can be pluralized with *-ob*. The clitic *=la* marks participant plurals. *=La* undergoes vowel assimilation when *-oñ* attaches to it in the exclusive (*laj+oñ* → *lojoñ*). Ch'ol pronouns are based off the Proto-Maya form *ha'- with vowel assimilation of the root to the vowel of the suffix. The forms *joñoñ* and *jatyety* are probably derived from **ja'-oñ* and **ja'-ety* (Hopkins et al. 2008). These forms are related to the *jiñ(i)* determiner in Ch'ol. Since the third person absolutive is null, its pronominal form is consequently *jiñ*.

While the set A markers are agreement markers (Coon 2013), I take the set B absolutive markers to be pronominal enclitics in Ch'ol as per Coon (2013:45). Following conventions in previous work (Coon 2013, Vázquez Álvarez 2011), I will still indicate absolutive markers with a hyphen '-', instead of '='.

Ch'ol has been described as having an inclusive and exclusive distinction (e.g. Kaufman & Justeson 2003, Law 2009, Coon 2010, Vázquez Álvarez 2011). In Ch'ol, the exclusive form is morphologically more complex than the inclusive and is formed by adding the first person marker *-oñ* to the inclusive form (Mora-Marín 2009:108). Effectively, the exclusive form contains the inclusive form. In §2.3, I will provide evidence that the inclusive is not a traditional inclusive form (speaker and hearer), but in fact a general first person plural form, unspecified for inclusion of hearer. The exclusive, on the other hand, is morphologically and semantically more specified to exclude the hearer.

2.2 Morphology of participant plurals in Ch'ol

Participant plurals in Ch'ol are marked with the =*la* clitic and set A or set B markers. The participant pluralizer can occur as enclitic on the predicate like in 1 where it appears on a verb (1a) or a noun (1b), marking first person inclusive.

(1) a. Tsa' majl-i-yoñ=**la**.
PRF go-IV-B1-**PART.PL**
'We left.'

b. k-otyoty=**la**
A1-house=**PART.PL**
'our house'

The =*la* clitic can also also appear as a proclitic on its host as in 2.

(2) **la=**k-otyoty
PART.PL=A1-house
'our house' (compare 1b)

Similarly, the exclusive marker, =*loj-oñ*, composed of the participant plural and the first person clitic, can also be an enclitic or a proclitic. It surfaces in its full form as an enclitic in 3.

(3) k-otyoty**=loj-oñ**
A1-house**=PART.PL-B1**
'our (excl) house'

As a proclitic, *=loj=oñ* surfaces as *loñ* in 4b from Vázquez Álvarez (2011:82). Its full form is not grammatical as a proclitic.

(4) a. **loñ=**k-otyoty
PL.EXCL=A1-house
'our houses'[2]

b. ***lojoñ=**k-otyoty
PL.EXCL=A1-house
'our houses'

When *=la* is an enclitic, it attaches after other second position enclitics. For instance in 5, *=la* attaches after the aspectual clitic *=tyo* 'still'.

(5) K-lumal=tyo**=la**.
A1-land=still**=PART.PL**
'It is still our town.' Adapted from Vázquez Álvarez (2011:81)

2.3 The inclusive as a general first person plural

As expected, the inclusive form is used in contexts where both speaker and hearer are referenced as in 6.

(6) Typical way for people to end a conversation with each other:
Mu'=tyo **la=k-**pejk-añ **la=k-**bä!
IMF=still PART.PL=A1-talk-NML PART.PL-A1-self
'We'll talk later!'

However, the inclusive plural has other extended usages. It is the default possessor when listing inalienable body-part terms in 7.

[2]This form and the one in 5 where the exclusive marker is a prefix is primarily used in the Tila dialect. Tumbalá dialect speakers use the form in 3 where the exclusive marker is a enclitic in possessive structures.

(7) a. *la=k-pix* = knee
b. *la=k-ok* = leg
c. *la=k-jol* = head

In 7, *lakpix* is not really 'our (yours and mine) inclusive knee', nor is *lakok* 'our (yours and mine) leg'. The first person inclusive is the general term used in this listing context.

Ch'ol speakers also use the first person inclusive in generic/impersonal contexts as in 8, much like English speakers use the second person.[3]

(8) Excerpt from instructions on how to build a house:
Mi **la=k**-ñajañ-ts'äp jiñi oy…
IMPF **PART.PL-A1**-first-lay.down DEM post
'First one (lit. 'we') puts in the posts…'

Finally, the inclusive form is used with many other nouns that do not literally mean that there is some relationship between speaker and hearer. Examples from respectful forms of address and some common nouns are given in Table 3. For respectful terms like *lakchich*, literally 'our older sister', the speaker can be directly addressing the listener and in this context it is clear that the general first person plural form does not include the addressee.

Table 3: Extended uses of the inclusive

Ch'ol	Literal translation	Usage/meaning
la-k-chich	our older sister	respectful form of address for elderly women
la-k-ña	our mother	elderly woman
la-k-tyaty	our father	elderly man
la-k-ty'añ	our language	the Ch'ol language
la-k-yum	our god	God

[3]This extended usage of the first person form in Ch'ol has interesting parallels with work on the Modern West Frisian impersonal pronoun *men*. Hoekstra (2010) provides evidence that the impersonal *men* can be used in contexts where the speaker is referring to herself. While the first person plural in Ch'ol is used in impersonal contexts, the Modern West Frisian impersonal form is used in first person contexts.

Based on the wider distribution of the inclusive, I conclude that the inclusive form is actually a general first person plural that does not have to make reference to a hearer.[4]

2.4 The exclusive as more specified

The exclusive, on the other hand, is only used in contexts explicitly excluding the hearer. In (9a), when the speaker is saying goodbye on behalf of a group, they use the exclusive form, which literally means 'we are leaving'. The inclusive form is infelicitous (9b).

(9) Context: As one group of people are leaving they say to another group (or person):

a. Sam-i-yoñ=ix**=loj-oñ**.
leave-IV-B1=already**=PART.PL-B1**
'Goodbye!' (Lit. We are leaving)

b. #Sam-i-yoñ=ix**=la**.
leave-IV-B1=already**=PART.PL**
'Goodbye!'

Similarly in an excerpt from naturalistic speech in 10, the speaker uses the exclusive form while relating an event that happened to her and another group of women. The addressee (the author of this paper) was not there when this event happened.

(10) Context: The speaker relating an event that happened to her and others, but not the addressee:
Che'=ta' ta' k-pijty-ä**=loj-oñ**.
PART=REA PRF A1-wait-TV**=PART.PL-B1**
'So, we (excl) waited for her.'

Finally, where a speaker is making a statement on behalf of a group about an addressee, the exclusive is used as in 11a. The inclusive plural is not used in this context as per 11b.

(11) a. Tyijikña-yoñ**=loj-oñ** che' wä-añ-ety=i
happy-B1**=PART.PL-B1** COMP here-EXT-B2=ENCL
'We are happy that you are here.'

[4]Throughout this paper, I will occasionally refer to the general first person plural form as inclusive, as it has traditionally been described as.

b. #Tyijikña-yoñ**=la** che' wä-añ-ety=i
happy-B1**=PART.PL** COMP here-EXT-B2=ENCL
'We are happy that you are here.'

In sum, the exclusive plural is both morphologically and semantically more complex. Morphologically, it is formed by adding the first person marker *-oñ* to the inclusive form. Semantically, it only occurs in contexts that explicitly exclude the hearer.

2.5 Summary

According to the actual usage of the inclusive form in Ch'ol, I recategorize it as a general first person plural for the purposes of this paper. Thus, the inclusive in Ch'ol does not necessarily include hearer.

Table 4: Summary of uses for first person plurals

Form	Traditional label	Usage	Recategorization
oñ=la	inclusive	generic contexts, speaker + hearer contexts, default possession with inalienable nouns, certain lexical items	General first person plural
oñ=loj-oñ	exclusive	To exclude hearer	Specified exclusive

3 Deriving Ch'ol first person plurals with binary features

3.1 A binary feature analysis for Ch'ol participant plurals

I propose a binary feature system with a combination of features from Nevins (2007), Harbour (2007),Watanabe (2013) and Bobaljik (2008) under a Distributed Morphology approach (Halle & Marantz 1993). Maximally attested person features are represented with the binary features [±speaker] and [±hearer] from Bobaljik (2008) in Table 5.

Table 5: Table adapted from Bobaljik (2008:206)

First person inclusive	[+speaker,+hearer]
First person exclusive	[+speaker,–hearer]
Second person	[–speaker,+hearer]
Third person	[–speaker,–hearer]

Recall again the generalization that =*la* occurs in all the participant plural forms. Based on the system in Table 5 alone, an additional feature is needed to capture =*la*: [±participant] (as in Nevins (2007); Harbour (2007)) given in Table 6.

Table 6: Table 5 updated with [±participant]

First person inclusive	[$\boxed{\text{+participant}}$,+speaker,+hearer]
First person exclusive	[$\boxed{\text{+participant}}$,+speaker,–hearer]
Second person	[$\boxed{\text{+participant}}$,–speaker,+hearer]
Third person	[–participant,–speaker,–hearer]

The updated feature system in Table 6 is advantageous for formalizing person, because [+participant] represents a natural class of speech act participants and can capture that =*la* shows up as a plural marker with persons that have a [+participant] value.

In order to capture the less specified first person plural, I posit a general first person that does not have a feature for [±hearer], given in Table 7.

Table 7: Table updated to reflect general first person

<u>General first person</u>	[+participant,+speaker]
First person exclusive	[+participant,+speaker,–hearer]
Second person	[+participant,–speaker,+hearer]
Third person	[–participant,–speaker,–hearer]

The analysis I propose for Ch'ol is featureally more complex than in Table 7, as the participant markers are composed of a person marker and the participant plural. Recall that set B markers *oñ* (first person) and *ety* (second person) are pronominal enclitics, and thus are taken to be full pronouns.[5] Therefore, I take them to be fully specified for features.

I posit the morphological spellouts of features in 12, with the addition of [±singular] for number.

(12) Ch'ol Vocabulary Items
  a. *-oñ* : [+participant,+speaker]
  b. *-ety* : [+participant,–speaker,+hearer]
  c. *=la* : [–singular] (in the context of [+participant])

I propose that the person (π) and number (#) nodes are separated, which is why number and person are expressed with different morphemes in Ch'ol.[6] The general first person plural has the features [+participant,+speaker] for [π] and [–singular] for [#]. The features [+participant,+speaker] are spelled out as *oñ* and, in the context of [+participant], [–singular] is spelled out as *=la*.

The exclusive form, *=loj-oñ* (*la* + *-oñ*), is derived from the general first person plural form by adding *-oñ*, the first person marker. However, adding *oñ* (with features [+participant,+speaker]) does not capture the semantic distribution of *=loj-oñ* as there needs to be a way to explicitly exclude the hearer. Therefore, I posit that to derive the exclusive, an additional [π] node is needed in the structure as the person marker *oñ* is repeated twice. I propose that to derive the exclusive from the inclusive, the feature values in the additional [π] node are [+participant,+speaker,**–hearer**]. The feature [–hearer] must be included in order to explicitly exclude the hearer, capturing the contexts in which the exclusive form, *=loj-oñ*, appears. However, no item in the vocabulary matches this set of features exactly. So, via the Subset Principle (e.g. Halle (1997)), given below, the phonological exponent that matches a *subset* of the features is inserted.

> Subset Principle Halle (1997): 'The phonological exponent of a Vocabulary item is inserted into a morpheme... if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not

[5] See also Woolford 1999, Preminger 2014, and Coon et al. 2014 for further arguments on the status of absolutive markers in Mayan languages as clitics.

[6] Preminger (2014) among others have proposed this for other Mayan languages.

> take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen.'

The item matching a subset of the features is *-oñ* in 12a, with the features [+participant,+speaker]. The second person marker *-ety* and participant pluralizer *=la* in 12c and 12b cannot be inserted as they have conflicting features. This analysis can therefore capture the morphology of the exclusive form and why first person *-oñ* appears in the exclusive form: it is the closest vocabulary item matching a subset of features needed to derive the exclusive form. The analysis is summarized in 13.

(13) Deriving the first person plural exclusive *oñ=loj-oñ*

Nodes:	[π]	[#]	[π]
Values:	[+participant,+speaker]	[–singular]	[+participant,+speaker,–hearer]
Vocab. insertion:	*-oñ*	*=la*	*-oñ* (via the Subset Principle)
Final form:	*-oñ*	*=loj*	*-oñ*

In sum, Ch'ol does not derive the inclusive/exclusive distinction in a way normally expected. There are no morphemes that independently spell out the distinction made in Table 5. Rather, the distinction is derived by adding the features [+participant,+speaker,–hearer] and via the Subset Principle *-oñ* is inserted to derive the exclusive from the general first person plural. This captures that the exclusive is both semantically and morphologically more complex.

This data also provides support for the necessity of a [±hearer] feature. It would be difficult to describe and explain the Ch'ol data without referring explicitly to an absence of a hearer in binary feature approaches like in Nevins (2007) or Harbour (2016) where the hearer feature is privative. Theories that do not have a [±hearer] feature would need to posit that two first person markers intensify, or emphasize, the first person and by doing so, exclude the second person. In other words, there would be more focus on first person and consequently a preference for its usage in an exclusive context. However, the analysis argued for here derives both the morphology and semantic distribution of the Ch'ol exclusive form without the need to stipulate anything about repeating the first person marker.

3.2 Choosing between the first person plurals

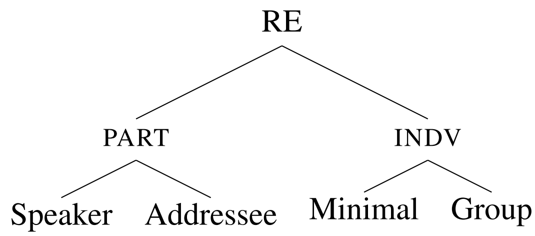
As shown above, the Ch'ol exclusive is more specified, with the inclusive/general first person plural used in other contexts. Speakers therefore choose the form that is most informative for the context. This can be derived through Heim (1991)'s MAXIMIZE PRESUPPOSITION, as has been done before for the choice of gender in Russian and Brazilian Portuguese (Bobaljik & Zocca 2011). The feature [–hearer] introduces a presupposition that the hearer is excluded. In a context explicitly excluding the hearer, the most informative form is *-oñ=loj-oñ*. In other contexts that do not explicitly exclude the hearer, like in generic instructions, forms of respect and other terms, speakers use the inclusive/general first person plural form. It is therefore important for this analysis that the inclusive form does not have a [hearer] feature and the exclusive form is specified for [–hearer].

4 A privative feature approach?

The above analysis assumes a binary feature approach with three binary features for person: [±participant], [±speaker], and [±hearer]. I argue that all these features must be binary in order to capture the Ch'ol data. This departs from some semi-binary approaches that do not have a [±hearer] feature (Nevins 2007, Harbour 2007). In this section, I discuss further how purely privative approaches do not seem to be able to account for the Ch'ol data. For reasons of space, I exemplify this with Harley & Ritter (2002)'s feature geometry; however, the issues this data poses for Harley & Ritter (2002)'s feature geometry also extend to Cowper & Hall (2005)'s feature geometry.

In the feature geometry for person and number in 14 from Harley & Ritter (2002), RE refers to referring expression (a pronoun, for instance), PART is the participant node under which the features [Speaker] and [Addressee] live, and [INDV] is the individuation node where features for number ([Group] and [Minimal]) are. Each feature is monovalent and only appears if it has a positive value.

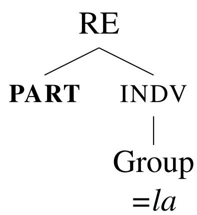
(14) Person and Number Features



There is a dependency relationship between features. Thus, [Speaker] implies the existence of [PART] and [Speaker] cannot be present when [PART] is not present.

To capture the Ch'ol data, =*la* would be the spellout of [Group] in the context participant (PART) in 15.

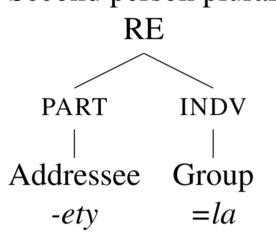
(15) When the PART node is activated, Group is spelled out as =*la*



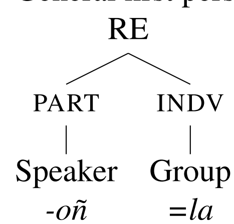
The analysis in 15 accounts for the distribution of =*la* in all the plural participant forms.

The second person plural -*ety*=*la* can be represented as in 16 and the general first person plural is the activation of [Speaker] and [Group] in 17.

(16) Second person plural



(17) General first person plural



The problem arises when trying to account for the exclusive form. To derive the morphology of the exclusive -*oñ*=*loj*-*oñ* with two speaker features, the formalism must have two speaker features activated as per 18. Like for the binary feature analysis proposed, person (π) and number (#/INDV) nodes are separated, with an extra node for person.

(18) Deriving the first person plural exclusive *oñ=loj-oñ*

Nodes:	[π]	[#/INDV]	[π]
Features:	PART | Speaker	Group	PART | Speaker
Vocab. insertion:	*-oñ*	*=la*	*-oñ*
Final form:	*-oñ*	*=loj*	*-oñ*

There needs to be a way to explicitly exclude the hearer from the featureal representation. However, a privative feature approach does not have a way of doing this. The absence of an [Addressee] node does not mean it is excluded from the meaning of the first person plural. Indeed English's 'we' pronoun is represented with the features in 17 but in some contexts it includes hearer and some it excludes the hearer. In 18, it is unclear how repeating two speaker features excludes the hearer: the absence of a feature does not mean it is necessarily excluded.

I conclude that the Ch'ol data necessitates an explicit exclusion of hearer in the exclusive form to account for its morphology and semantic distribution.

5 Implications & conclusion

In this paper, I recategorized the Ch'ol inclusive/exclusive distinction as a general first person plural and a specific first person exclusive. I have proposed an analysis for Ch'ol person features in a binary feature system. The exclusive form is derived from the general first person plural by adding [+participant,+speaker,–hearer] in an additional person node. No item in the Vocabulary matches these features exactly so via the Subset Principle, *oñ*, which contains a subset of the features ([+participant,+speaker]), is inserted. Speakers choose the form that is most informative for the situation. This can be accounted for by using Heim (1991)'s Maximize Presupposition.

This analysis has implications for theories of person. The binary feature [±hearer] departs from recent work on binary features that posits a privative hearer feature (Nevins 2007, Harbour 2007, 2016). Feature geometries, such as the one proposed by Harley & Ritter (2002), do not have a way to explicitly exclude the hearer. I argue that in order to account for the more complex semantics and morphology of the exclusive in Ch'ol, there must be a binary hearer feature.

Recent research has also argued for the existence of a binary hearer feature (Watanabe 2013, Despić & Murray 2018). For instance, Watanabe (2013) gives evidence for [–hearer] in Fula (Niger-Congo) based on the split behavior of person morphemes with [+speaker] and [–hearer] values. In Fula, when the [–hearer] feature is necessary, like in first person plural exclusive and third person, the person marker is placed before the verb. Otherwise, the person marker comes after the verb. Though there is not a morpheme in Fula that spells out [–hearer], its presence is necessary in order to account for the distribution of morphemes, i.e. if they are prefixes or suffixes. The binary feature analysis argued for here seems to better capture the variety seen across the world's languages.

REFERENCES

BOBALJIK, JONATHAN DAVID. 2008. Missing persons: A case study in morphological universals. *The Linguistic Review* 25.203–230.

BOBALJIK, JONATHAN DAVID, and CYNTHIA LEVART ZOCCA. 2011. Gender markedness: the anatomy of a counter-example. *Morphology* 21.141–166.

COON, JESSICA. 2010. *Complementation in Chol (Mayan): A theory of split ergativity*. Cambridge, MA: MIT dissertation.

COON, JESSICA. 2013. *Aspects of split ergativity*. Oxford University Press.

COON, JESSICA; PEDRO MATEO PEDRO; and OMER PREMINGER. 2014. The role of case in A-bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation* 14.179–242.

COWPER, ELIZABETH, and DANIEL CURRIE HALL. 2005. The pieces of pi. *Proceedings of the 2004 Canadian Linguistic Association annual conference*, ed. by Marie-Odile Junker, Martha McGinnis, and Yves Roberge.

DESPIĆ, MILOJE, and SARAH MURRAY. 2018. On binary features and disagreeing natural classes: Evidence from Cheyenne and Serbian. *The Linguistic Review* 35.219–241.

HALLE, MORRIS. 1997. Distributed morphology: Impoverishment and fission. *MITWPL 30: Papers at the interface*, ed. by Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, 425–449. Cambridge, MA: MITWPL.

HALLE, MORRIS, and ALEC MARANTZ. 1993. Distributed Morphology and the pieces of inflection. *The view from building 20*, ed. by Kenneth Hale and Samuel J. Keyser. Cambridge: MIT Press.

HARBOUR, DANIEL. 2007. *Morphosemantic number: From Kiowa noun classes to UG number features*. Dordrecht: Springer.

HARBOUR, DANIEL. 2016. *Impossible persons*, vol. 74. MIT Press.

HARLEY, H., and E. RITTER. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78.482–526.

HEIM, IRENE. 1991. Artikel und Definitheit. *Semantik: ein internationales Handbuch der Zeitgenössischen Forschung*, 487–535.

HOEKSTRA, JARICH. 2010. On the impersonal pronoun men in Modern West Frisian. *The Journal of Comparative Germanic Linguistics* 13.31–59.

HOPKINS, NICHOLAS A; AUSENCIO CRUZ GUZMÁN; and J KATHRYN JOSSERAND. 2008. A Chol (Mayan) vocabulary from 1789. *International Journal of American Linguistics* 74.83–114.

KAUFMAN, TERRENCE, and JOHN S JUSTESON. 2003. *A preliminary Mayan etymological dictionary*. FAMSI.

LAW, DANNY. 2009. Pronominal borrowing among the Maya. *Diachronica* 26.214–252.

MORA-MARÍN, DAVID F. 2009. Reconstruction of the Proto-Ch'olan demonstrative pronouns, deictic enclitics, and definite articles. *Transactions of the Philological Society* 107.98–129.

NEVINS, ANDREW. 2007. The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25.273–313.

NOYER, ROBERT ROLF. 1992. *Features, positions and affixes in autonomous morphological structure*. Massachusetts Institute of Technology PhD thesis.

PREMINGER, OMER. 2014. *Agreement and its failures*, vol. 68. MIT Press.

VÁZQUEZ ÁLVAREZ, JUAN J. 2011. *A grammar of Chol, a Mayan language*. Austin, TX: University of Texas Austin dissertation.

WATANABE, AKIRA. 2013. Person-number interaction: Impoverishment and natural classes. *Linguistic inquiry* 44.469–492.

WOOLFORD, ELLEN. 1999. Ergative agreement systems. *Maryland Mayfest, University of Maryland*.

The attributivizing marker in Andi (Nakh-Daghestanian): clitic, affix, or both?

TIMUR MAISAK

Institute of Linguistics, Russian Academy of Sciences /

National Research University Higher School of Economics[*]

1 Introduction

Andi, a Nakh-Daghestanian language spoken in the Republic of Daghestan, Russia, possesses a grammatical marker (*=sːi*) with the ATTRIBUTIVIZING function, which makes an attribute from a constituent which cannot be an attribute by itself. Thus, in 1, the locative noun phrase ***rešinola*** 'in the sky' (***rešin*** 'sky' in the superessive case), being attributivized by means of *=sːi*, thereby acquires the ability to modify the head noun ***šan*** 'spirit'. In 2, *=sːi* attributivizes a verb phrase ***c'ek'irodi c'ek'irdaqi hit'omado*** '(one) kid is saying to (other) kids' headed by a finite present progressive form: with the attributivizer, this phrase becomes a modifier of the head noun ***dan*** 'thing'.[1]

(1) bihu-rihi urʁunni-dːu rešin-o-la=**sːi** šan.

much-TEMP think-PRF [[sky-OBL-SUPER]=ATTR spirit.ABS]

'The heavenly (lit. which is on the sky) spirit thought for a long time.' (Tales)

---

[*] This research is part of the project 'Andi morphosyntax in a typological perspective' supported by the Russian Foundation for Basic Research (RFBR), grant No. 18-012-00852a. My thanks are due to Chopan Gaziev for his native speaker judgements on Andi, and to Steven Kaye for the help with the English of the paper.

[1] In the examples to follow, the constituent which gets attributivized, as well as the whole noun phrase with its head and modifier, are marked with square brackets in the glossing line. Abbreviations: ABS – absolutive; AD – localization 'near' (1); ADD – additive; AFF – affective (case); ANTE – localization 'in front'; AOR – aorist; APUD – localization 'near' (2); ATTR – attributivizer; CAUS – causative; COMIT – comitative; COMPAR – comparative; CONT – localization 'in contact'; COP – copula; DAT – dative; DEM – demonstrative; DEP – subordinator; ELAT – elative; ERG – ergative; F – feminine; F/N – feminine or neutral; FUT – future; GEN – genitive; HAB – habitual; III – third (neutral) gender; IMP – imperative; INDEF – indefiniteness; INCL – inclusive; INT – intensive; INTER – localization 'in a mass'; INTJ – interjection; IV – fourth (neutral) gender; LAT – lative; M – masculine; NEG – negation; OBL – oblique stem; PFV.PTCP – perfective participle; PL – plural; PRET – preterite; PRF – perfect; PROGR – progressive; PRS – present; Q – polar question; QUOT – quotative; SEQ – sequential; SUPER – localization 'on top'; TEMP – temporal; V – fifth (neutral) gender; WH – content question; 3 – third person.

(2)

boc'u-‹b›o	onš:igu	anɬi-d:u	c'ek'ir-o-di	c'ek'ir-da-qi
wolf.OBL-AFF<III>	once	hear-PRF	[[kid-OBL-ERG	kid-PL.OBL-APUD
hit'om-ado=**s:i**	dan.			
say-PROGR]=ATTR	thing.ABS]			

'The Wolf once heard what (lit. the thing which) one kid said to the other kids.' (Tales)

Those word classes and word forms which can be used as attributes by themselves (e.g. adjectives, demonstratives, numerals, participles, and the genitive case) do not require the use of the attributivizer *=s:i*, and in fact they do not allow it.

In this paper, I will address the question of the morphosyntactic status of the attributivizer *=s:i*. As I show in section 2, this marker possesses certain properties which make it similar to a clitic (in particular, an enclitic), that is a morpheme with a freer syntactic distribution than is expected of affixes. On the other hand, though, the behaviour of the attributivizer is more like that of an affix in other respects, as will be discussed in section 3. In section 4 I summarize the morphosyntactic properties of *=s:i*, suggesting potential ways of analysing this marker and looking at its cross-linguistic parallels. Section 5 concludes the paper.

Andi belongs to the Andic group of the Avar-Andic branch of the Nakh-Daghestanian (also known as East Caucasian) language family. Closely related languages include Akhvakh, Bagwalal, Botlikh, Chamalal, Godoberi, Karata, and Tindi, all of which are unwritten minority languages spoken in the north-western part of Daghestan by a few hundred to a few thousand speakers each. Avar, the closest relative of the Andic languages, is, in contrast, one of the major languages of Daghestan, with more than 700,000 speakers and an elaborate dialectal system. Like the languages of other Nakh-Daghestanian branches,[2] those in Avar-Andic are morphologically ergative, mostly agglutinative, left-branching, SOV languages with rich case systems (including many locative forms) and gender agreement (comprising three to five or six genders).

Andi, which is spoken in a few villages of Botlikhsky District of Daghestan, is not very well documented. Although a number of grammatical sketches have been produced since the

[2] The other branches of Nakh-Daghestanian are Nakh, Tsezic, Lak, Dargwa, Lezgic and Khinalug.

early 20th c. (see Dirr 1906, Sulejmanov 1957, Cercvadze 1965, and Salimov 2010, based on the author's 1968 dissertation), all these sources provide only general overviews of phonology and morphology of the language. The data on which the present paper is based mostly come from texts and, to a lesser degree, from elicitation (my own fieldwork in Daghestan in 2015–2017). The available text corpus of Andi is not very big and includes, besides a number of texts appended to the four aforementioned grammars, two books published in the 2010s which represent the first serious examples of literary production in Andi. The first of these is a collection of a dozen world fairy-tales ('Tales' in the examples below), and the second is the translation of the Gospel of Luke (henceforth, 'Luke').

2 The attributivizing marker as a clitic

The main reason it is tempting to treat the attributivizer *=sːi* as a clitic is its low selectivity and phrase-level distribution. While affixes (like case suffixes on nouns or tense-aspect suffixes on verbs) combine with stems of particular word classes, the attributivizer modifiers whole phrases (making non-attributive phrases attributive). It is thus hosted not by stems, but by fully inflected words of various word classes. In particular, *=sːi* can be hosted by case forms of nouns, by adverbs and postpositional phrases, and by finite verb forms. Given that in Andi phrases are head-final, the attributivizer is always hosted by the head which at the same time happens to be the last word of the phrase.[3]

In section 1, we have already seen an example of locative noun phrase attributivization. Sentence 3 provides another instance of this kind: the phrase *ʁʷannar honʟikːu* 'from the Andi village', whose head (bearing interelative case inflection) hosts the attributivizer, becomes the modifier of *akademik* 'academician' (lit. 'the academician who is from the Andi village'). Example 4 illustrates the attributivization of the dative noun phrase (in benefactive function) *χːuntʼuʟi bošilloj* 'for a piglet'. In 5, the postpositional phrase *iʟičʼu sedu* 'in front of us' becomes an attribute of *hon* 'village'.

[3] This means that in case of the attributivizer we cannot distinguish between head marking and edge marking (the word hosting *=sːi* is the head and at the same time the rightmost word of the phrase). This is not of crucial importance from the point of view of the clitic/affix distinction, though, as alongside head inflectional marking, edge inflection is also attested in the languages of the world (see Spencer & Luís 2012:126–132, 293ff. for discussion).

(3)

ɢʷanna-r	hon-ʟi-k:u=**s:i**	akademik	χizri	amirχanov
[[Andi-V	village-INTER-ELAT]=ATTR	academician.ABS]	Khizri	Amirkhanov

'Academician from the Andi village, Khizri Amirkhanov'

(4)

ho-b	e‹b›i=ʁil,	χ:unt'.u-ʟi	boši.llo-j=**s:i**	milki=le,
DEM-IV	what<IV>=WH	[[pig-GEN	baby-DAT]=ATTR	house.ABS]=Q
huns:u	qala=le?			
not	fortress.ABS=Q			

'What is it, a house for a piglet, or a fortress?' (Tales)

(5)

hinidi-r	iʟi-č'u	sedu=**si**[4]	hon-ʟi	w-olo-m-ul
DEM-V	[[we.INCL-CONT	before]=ATTR	village-INTER]	M-PL.go-IMP-PL
bis:il.				
you.PL.ABS				

'Go to the village which is ahead of us.' (Luke 19:30)

Quite remarkably, the attributivizer *=s:i* can be hosted by a constituent which already hosts another enclitic. This case is illustrated in 6–9 below. The comitative marker *=logu* is close to being a case marker, though unlike the other case markers it does not follow the oblique stem of a noun but is hosted by the absolutive noun phrase. In 6, it is the comitative phrase *ɢuwwati=logu* 'with strength' that is attributivized and modifies the head noun 'lion'. The comparative marker *=ga* can modify noun phrases (in various cases), as well as adjectives and participles. In 7, *=ga* modifies the participial clause 'which happened in the days of Noah'; the comparative phrase

[4] In examples from the Gospel of Luke translation, we keep the transliteration close to the (Cyrillic) original. For some reason, the intensive /s:/ is rendered inconsistently in this source, often appearing as single <c> /s/ instead of double <cc> /s:/.

'similar to the one which happened in the days of Noah' undergoes attributivization and modifies the head noun *dan* 'thing'.

(6) b-ik'ʷo-dːu inna=bolo zolo ɢuwwati=**logu=sːi** ʁalbac'i

III-be-PRF when=INDEF [very [strength.ABS=COMIT]=ATTR lion.ABS]

'One time, there lived a strong lion (lit. a lion being with strength).' (Tales)

(7) adam-ʟi moč'i-šu-b zamana-l-la=lo b-uʟi-dja

people-GEN child-OBL.M-GEN(IV) time-OBL-SUPER=ADD IV-happen-FUT

nuħi awaregi-šu-b zamana-l-la

[[Noah.ABS prophet-OBL.M-GEN(IV) time-OBL-SUPER

b-uʟi-b=**ga=si** dan.

IV-happen-PFV.PTCP=COMPAR]=ATTR thing.ABS]

'In the days of the Son of Man it will be just as it was (lit. the thing like the one that happened) in the days of Noah.' (Luke 17:26)

The situation is even more interesting with the quotative marker *=ʁodːu*, which is a typical right-edge enclitic in that it occurs on the rightmost word of a quotation. In particular, person or place names are usually introduced with the help of the quotative marker. In 8, the girl's name *mašenʲka* is followed by the quotative, and the whole combination is attributivized to modify the noun phrase 'granddaughter' (= daughter's daughter). The next example 9 comprises the sentential quote 'Physician, heal yourself!', which is also marked with a quotative[5] and is also attributivized, thus becoming an attribute of *kici* 'proverb'. The noun phrase with the attributivized quotation can be rendered literally as 'the *Physician, heal yourself!*-saying proverb'.

[5] The quotative marker *=ʁodːu* is prosodically dependent on its host. Nonetheless, in written texts it tends to be separated from the quoted material by means of a dash (cf. 9), which makes it look an independent word.

(8)

j-ik’ʷo-d:u	hege-l-u-j	mašenʲka=**ʁod:u=s:i**	jošu-LI
F-be-PRF	DEM-PL-OBL-GEN(F)	[[Mashenka.ABS=QUOT]=ATTR	daughter.OBL-GEN
joši=lo			
daughter.ABS]=ADD			

‘And they had a granddaughter whose name was Mashenka.’ (Tales)

(9)

haj-haj,	«toχtur,	lesedu-‹w›a	men-ni	du-j=gu	daru
INTJ	[[doctor.ABS	before-LAT<M>	you-ERG	you-DAT=INT	medicine.ABS
ih-o», –	**ʁod:u=si**	kici	šam-ja	bis:i-di	di-j.
do-IMP	QUOT]=ATTR	proverb.ABS]	throw-FUT	you.PL-ERG	I-DAT

{Jesus said to them,} ‘Oh, you will now quote to me the proverb “Physician, heal yourself!”’ (Luke 4:23)

The occurence of *=s:i* on constituents already hosting clitics is a strong argument in favour of the clitic status of the attributivizer itself: of course, we do not expect affixes to appear after clitics, but clitic-clitic sequences are perfectly normal. Zwicky & Pullum (1983:504) formulate this as one of their criteria for the clitic vs. affix distinction, viz.: ‘Clitics can attach to material already containing clitics, but affixes cannot’.

3 The attributivizing marker as an affix

In Andi, attributes such as adjectives, participles, or genitives inflect when they agree with their head for plural number, or when they head noun phrases themselves. Thus, whereas with a head in the singular the adjective *w-oč’uχa* ‘big’ occurs in the ‘unmarked’ form (see 10a),[6] it

[6] Some Andi adjectives (as well as some verbs, adverbs, postpositions and even case markers) include an affixal gender agreement slot: in the case of *w-oč’uχa*, this is a prefix. Within noun phrases, attributes with gender agreement slots agree in gender with their head noun.

takes the plural suffix -*ol* when agreeing with a plural head (see 10b). Headless adjectives inflect for case and number: in the absolutive, the form is again 'unmarked', whereas all other case forms are derived from an oblique stem marker, which depends on the gender of the referent. When the referent of a noun phrase is masculine human, the oblique stem marker is -*š*-, and otherwise it is -*l*-. Case suffixes follow these oblique stem markers, cf. the ergative case in -*di* in 10c-d.

(10) a. w-oč'uχa wošo

M-big boy.ABS

'a big boy'

b. w-oč'uχ-ol wošo-wl

M-big-PL boy-PL.ABS

'big boys'

c. w-oč'uχa-š-di

M-big-OBL(M)-ERG

'a big one, male <e.g. said>'

d. j-eč'uχa-l-di

F-big-OBL(F/N)-ERG

'a big one, female <e.g. said>'

A notable characteristic of phrases attributivized by means of the marker *=s:i* is that in the same morphosyntactic contexts they undergo the same affixal inflection as the other attributes mentioned above. In particular, when an attributivized phrase with *=s:i* agrees in the plural with its head, the plural suffix -*ol* is used, which fuses with the attributive marker (**=s:i* + -*ol* > -*s:ol*). The use of plural attributivized phrases is illustrated in 11 and 12 below. In 11, the attributive phrase '(one which is) in the Andi language' takes plural inflection, because it agrees with the plural head *muχal* 'fairy-tales'. In 12, the attributivized comitative phrase '(one which is) with honey' agrees with the plural head *bak'ʷolol* 'skins'.

(11) dunja-l-la baχunni-r=s:i χalɢi-l-ʟ-ol ɢʷanna-b

world-OBL-SUPER live-PROGR=ATTR people-OBL-GEN-PL [[Andi-IV

mic':i-l-la=**s:-ol** muχa-l

tongue-OBL-SUPER]=ATTR-PL tale-PL.ABS]

{Book title:} 'World folk-tales in Andi (lit. tales in the Andi language of the peoples living in the world)' (Tales)

(12) tiχolɬi din per-dul=lo inkʷa-l-č'u=reχu,

harness.CAUS.AOR I.ERG bee-PL.ABS=ADD cart-OBL-CONT=SEQ

b-ok'ʷolɬi hege-l-i-la hunc':i=logu=**s:-ol**

IV-PL.be.CAUS.AOR DEM-PL-OBL-SUPER [[honey.ABS=COMIT]=ATTR-PL

bak'ʷol-ol=lo=reχu, w-ulon išʷa.

skin-PL.ABS]=ADD=SEQ M-go.AOR home

'I harnessed the bees to the cart, loaded animal skins with honey on them, and went home.' (Tales)

Headless attributivized phrases take case morphology: in oblique cases, case suffixes follow the oblique stem markers (*-š-* in the masculine, *-l-* in the feminine/neuter), while in the absolutive the form in *=s:i* remains unmarked. For example, in 13, the headless attributivized phrase '(one) with a pale face' takes the apudessive, which marks the addressee of speech. In 14, the headless attributivized phrase '(one who) goes hunting' takes the genitive case marking the adnominal possessor.

(13) bosotannib riži=logu=**s:i-š:-qi** hit'on-d:u hege-l-d:i...

[[pale face.ABS=COMIT]=ATTR-OBL(M)-APUD] say-PRF DEM-OBL(F/N)-ERG

'She said to the one with a pale face...' (Tales)

(14) se-w emi=bolo c':adi-b čoniwo-di

one-M who=INDEF drink-PFV.PTCP [hunting-LAT

w-ulin-no=**s:i-š:u-b** k'otu-di...

M-go-HAB]=ATTR-OBL(M)-GEN(III)] horse.OBL-ERG

'A horse of someone drunk who was going hunting...' (Tales)

Note that when a noun phrase is attributivized and the attributivized phrase becomes the head of a noun phrase, there are two case/number slots in a word hosting the attributive marker, one 'internal' (case/number of the phrase which was an input to attributivization) and one 'external' (case/number of the attributivized phrase). For example, in 15, *ʁuluqillasišχo* is the result of attributivization of the locative noun phrase *ʁuluqilla* 'on service' (< *ʁuluqi* 'service' in the superessive case), and the resulting headless phrase *ʁuluqilla=si* 'one who is on service' takes the adlative case which here marks the prospective possessor.

(15) r-eχud-‹r›a ič:i ʁuluqi-l-la=**si-š-χo** χ:uča...

IV-back-LAT<IV> give.AOR [[service-OBL-SUPER]=ATTR-OBL(M)-AD.LAT] book.ABS

{Then Jesus rolled up the scroll and} 'gave it back to the attendant...' (Luke 4:20)

The problem which arises given the inflectional behaviour of attributivized phrases is the following: if *=s:i* is really a clitic (specifically one which can follow other clitics), how can it be followed in turn by clearly inflectional affixes? Is the attributivizer really a clitic at all?

4 The attributivizer in Andi: clitic or affix?

Like WORD, one of the central concepts in linguistics (cf. Haspelmath 2011), the concept of CLITIC is well known to be problematic. The reason for this is that although in a particular language it may be clear what language-specific criteria set bound affixes apart from less bound (clitic-like) markers, it seems to be impossible to suggest a consistent list of universally applicable criteria defining clitics cross-linguistically. Even Spencer & Luís (2012b:xiii), in the foreword to their book on clitics, confess: 'There's an important sense in which we don't actually believe in the existence of clitics' (while at the same time they admit that 'it's worthwhile to study in great detail the kinds of properties that have been ascribed to clitics'). In another influential paper, Spencer & Luís (2012a) approach the definition of a CANONICAL CLITIC as lying at the intersection of two dimensions: whereas the formal properties of a canonical clitic are those of a CANONICAL AFFIX, its distributional properties are those of a CANONICAL FUNCTION WORD, see 16 for a list of canonical clitic properties.

(16) Canonical properties for clitics (Spencer & Luís 2012a:130, 134)

a. Canonical form properties for clitics

(i) A clitic consists of a monomoraic CV syllable

(ii) A clitic is prosodically unspecified and hence is prosodically dependent on some other adjacent element

b. Canonical distributional properties for clitics

(i) A clitic is placed with respect to the syntactic phrase bearing the functional property expressed by the clitic (its morphosyntactic domain)

(ii) A clitic canonically takes wide scope over a coordinated phrase with which it is in construction

The morphosyntactic behaviour of the attributivizer *=sːi* comes close to that of the canonical clitic in Spencer & Luís's approach. Like typical affixes, *=sːi* is monosyllabic and lacks prosodic prominence, it requires a host and cannot occur in isolation (e.g. as an independent answer to a question), nor can it be focused. However, unlike affixes, the attributivizer has phrase-level function (turning non-attributes into attributes), it is hosted by already inflected words and shows low selectivity towards its host (displaying what Spencer & Luís call 'promiscuous attachment'). Moreover, *=s:i* does not have idiosyncratic relations with any of its hosts, keeping the same form and function irrespective of the particular host involved. The only criterion of canonical clitics that *=s:i* does not meet is wide scope over a coordinated phrase:[7] a coordinated phrase cannot be attributivized, although two attributivized phrases can be coordinated.

(17) woc:u-łu=**s:i**=lo jec:u-łu=**s:i**=lo χ:uča

[[brother.OBL-DAT]=ATTR=ADD [sister.OBL-DAT]=ATTR=ADD book.ABS]

'the book for the brother and sister'

However, although the Andi attributivizer *=s:i* looks very much like a canonical clitic, the clitic analysis faces the following problem: when the attributivized phrases need case/number marking, the word hosting the attributivizer takes its case/number inflections on top of *=s:i*! Thus, even if the attributivizer is preceded by another clitic itself, the case/number affixes follow, so what we have in examples 12–15 above is affixation after cliticization.

Interestingly, although in other Andic languages there are attributive markers functionally close (and cognate) to the Andi *=s:i*, not every aspect of behaviour of the Andi attributivizer is true of them. To judge from the existing descriptions (which, unfortunately, are not always detailed as far as the use of attributivizers is concerned), Bagwalal *=s:*, Chamalal *=is:*, Godoberi *=s:u* and a few other such markers are also clitic-like in showing 'promiscuous attachment'. Thus, the Bagwalal attributivizer can be hosted by noun phrases in locative cases, by postpositional phrases and adverbs (Kibrik et al. 2001:175); however, no examples are provided which show further case/number inflection on top of the attributivizer. For Godoberi, the latter possibility is explicitly rejected in Kibrik (1996:27–28): *=s:u* is described as the marker of 'adjective derivation', attaching to nouns, adverbs and postpositions, but it is mentioned that all such derivatives 'may occur only adnominally. No instance of substantivization of such attributes was found'.

[7] The means of coordination in Andi is the additive enclitic *=lo*, which attaches symmetrically to both conjuncts. With respect to the attributivizer the additive can take an external position only, and cannot occur within the attributivizer's scope.

At the same time, in other branches of Nakh-Daghestanian there are languages with attributivizing morphemes which allow further inflection of attributivized phrases. Such are, for example, the attributivizer *-se* in Tanti (Dargwa branch) and the attributive *-tːu* (*-du*) in Archi (Lezgic branch), see Sumbatova & Lander (2014:191–200) on the former and Kibrik (1977:113–116) and Bond & Chumakina (2016:48–53) on the latter. As the two following examples show, headless attributivized phrases take case/number inflection in Tanti Dargwa and Archi as they do in Andi.

(18) Tanti Dargwa (Sumbarova & Lander 2014:212)

buretːa.li-cːele-se.li-sa-r	w-ibšː-ib
[[axe.OBL-COMIT]-ATTR.OBL-ANTE-ELAT]	M-run.away:PFV-PRET

'[He] ran away from the one with an axe'.

(19) Archi (Bond & Chumakina 2016:52)

doš.mi-s-du-m-mi-s	χarak	is	i
[[sister.OBL-DAT]-ATTR-IV-OBL-DAT]	behind	IV.I.GEN	IV.be.PRS

'Mine [my chest for dowry] is behind the one that is for my sister.'

In both the Tanti Dargwa and the Archi descriptions, the authors treat the respective attributivizing morphemes as suffixes, not clitics, although there is no discussion of this choice. In any case, both Tanti Dargwa *-se* and Archi *-tːu* (*-du*) face the same problem of mixed affix/clitic status. Which is more important in their analysis, their clitic-like promiscuous attachment or their affix-like position in the synthetic word form after the (phrasal) 'base' and before other inflections? Restricting myself here to the Andi attributivizer, I suggest that there seem to be two reasonable ways to treat morphemes with such a free distribution.

First, the Andi *=sːi* can be analysed as an INFLECTING CLITIC, that is a clitic which alongside the unmarked form *=sːi* (used in the attributive position and in the absolutive singular) has a whole range of case/number forms, in particular *=sː-ol* (plural), *=sːi-šːu-b* (genitive), *=sːi-šː-qi* (apudessive), *=si-š-χo* (adlative) and so on. Under this approach, the (heads of) attributivized phrases do not themselves inflect, but rather the phrases host attributivizing clitics already inflected for the necessary case/number values. Thus, in 20 (a modified version of 11) it is a ‘plural attributive’ *=s:ol* that is attached to the phrase *ɢ^{w}annab mic’:illa*, and in 21 (a modified version of 13) it is an ‘apudessive, masculine attributive’ *=s:iš:qi* that is attached to the phrase *bosotannib rižilogu*.

(20) ɢ^{w}anna-b mic’:i-l-la **=s:-ol** muχa-l

[[Andi-IV tongue-OBL-SUPER] =[ATTR-PL]] tale-PL.ABS

‘folk-tales (which are) in Andi’

(21) bosotannib riži=logu **=s:i-š:-qi**

[[pale face.ABS=COMIT] =[ATTR-OBL(M)-APUD]]

‘(said) to the one with a pale face’

As a partial cross-linguistic parallel for an inflecting attributivizer/nominalizer, cf. one of the enclitic postpositions (or enclitic case markers) in Hindi (see Thakur 1997, Spencer 2005). The genitive enclitic *k-*, hosted by the possessor, agrees in number/gender with the head (the possessee), so there is a ‘masculine singular genitive’ clitic, a ‘feminine singular genitive’ clitic, etc.

(22) Hindi (Thakur 1997)

a. raam **k-aa** ghar

[Raam GEN-3.M.SG] house(3.M.SG)

‘Raam’s house’

b. raam **k-ii** kitaab

[Raam GEN-3.F.SG] book(3.F.SG)

‘Raam’s book’

For Hindi, however, the analysis of 'genitive postpositions' as agreeing is absolutely natural: historically, they go back to participles of auxiliary verbs (in particular, 'do'), so in their agreement they keep the inflectional possibilities of their sources (cf. Payne 1995). In Andi, there is no evidence that, for example, *=s:-ol* in 11 or *=s:i-š:-qi* in 13 go back to inflecting content words. To the contrary, it seems that the attributivizer is quite ancient and its origin should be traced back not even merely to Proto-Andic, but to a deeper level. The attributivizing morphemes *-s:a* in Lak and *-si* / *-se* / *-ze* / *-ci* in Dargwa varieties are most probably cognates of the Andic *-s:* / *-s:i* / *-s:u* / *-is:* (cf. Alekseev 2003:164) and may have enjoyed their affix/clitic status for millennia. The historical source of these markers (this marker?) might have been entirely uninflectable, meaning that case/number inflection in (some of) the modern languages may represent a late secondary development.

Under the second approach, the Andi *=s:i* should rather be recognized as a (non-canonical) derivational affix that turns a non-attributive phrase into an attributive one, making its host a kind of an 'attributive stem' irrespective of its morphological structure. That is, when a phrase undergoes attributivization, it – or at least its head, which hosts the attributivizer – becomes reinterpreted as a bound stem, which takes on further inflection according to the attributive inflection type (illustrated above in 15). At this stage of derivation, it does not matter whether the original phrase bears case/number morphology itself, or hosts other clitics. Thus, when the superessive noun phrase *ɢʷannab mic'ːilla* 'in the Andi language' become attributivized, *ɢʷannab mic'ːillas:i-* turns into an attributive stem meaning 'one which is in the Andi language', which takes a plural marker when agreeing with the plural head 'tales' (23). Likewise, *=s:i* attributivizes the comitative phrase *bosotannib riži=logu* 'with a pale face', which becomes an attributive stem *bosotannib rižilogus:i-* 'one that is with a pale face', so that the stem can take the apudessive case in the argument position (24).

(23)	**ɢʷanna-b**	**mic'ːi.l-la-s:**-ol	muχa-l
	[[[Andi-IV	tongue.OBL-SUPER]-ATTR]-PL]	tale-PL.ABS
	'folk-tales (which are) in Andi'		

(24) **bosotannib** **riži-logu-s:i-**š:-qi

[[[[pale face.ABS]-COMIT]-ATTR]-OBL(M)-APUD]

'(said) to the one with a pale face'

5 Conclusion

It has long been acknowledged that the affix/clitic distinction is fuzzy, and that language-specific markers may correspond to many various points on the affix–clitic continuum, depending on their phonological, prosodic, morphological and syntactic properties. The attributivizer *=s:i* in Andi (and, seemingly, similar morphemes in at least some other Nakh-Daghestanian languages) is definitely one such element with intermediate status. Although very similar to a clitic with low host selectivity, *=s:i* is at the same time a kind of a derivational affix in the string of affixes making up a word form.

Still, even in this situation, different analytical choices are possible in the treatment of the attributivizer. It can be analysed as a clitic which is close to an inflectable function word, if we decide that it is the attributive marker in its own right that bears case/number inflection. Alternatively, *=s:i* can be conceived as a clitic which is closer to an affix, one which syntactically turns non-attributive phrases into attributive ones and morphologically renders such phrases, in effect, attributive 'stems' – stems which inflect further like adjectives. Under the second approach, the strict affix vs. clitic qualification of *=s:i* indeed does not make much sense: the attributivizer behaves as a derivational suffix, whose distribution is at the same time much freer than expected for affixes of that kind. This is, however, perfectly in line with the understanding of clitics as 'phrasal affixes' in some recent approaches, that is as 'overt morphological markers of the morphosyntactic properties of phrases' (Anderson 2005:83). In other words, a phrasal affix is 'a morphological object (an affix) that happens to be attached to whatever happens to come at the edge of a phrase' (Spencer & Luís 2012b:107, also 126). Mutatis mutandis, the Andi *=s:i* can be approached as a phrasal affix with attributivizing function which attaches to the last word (= head) of a phrase, displaying a certain freedom in the choice of phrases which it can turn into attributives.

REFERENCES

ALEKSEEV, M. E. 2003. *Sravnitel´no-istoričeskaja morfologija naxsko-dagestanskix jazykov. Kategorii imeni* [A comparative-historical morphology of the Nakh-Daghestanian languages: Nominal categories]. Moscow: Academia.

ANDERSON, STEPHEN R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.

BOND, OLIVER, and MARINA CHUMAKINA. 2016. Agreement domains and targets. *Archi: Complexities of agreement in cross-theoretical perspective*, ed. by Oliver Bond, Greville G. Corbett, Marina Chumakina, and Dunstan Brown, 43–76. Oxford: Oxford University Press.

CERCVADZE, I. I. 1965. *Andijskij jazyk* [The Andi language]. Tbilisi: Mecniereba.

DIRR, A. M. 1906. Kratkij grammatičeskij očerk andijskogo jazyka [A short grammatical sketch of Andi]. *Sbornik materialov dlja opisanija mestnostej i plemën Kavkaza* 36. Tbilisi.

HASPELMATH, MARTIN. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1).31–80.

KIBRIK, A. E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka. T. 2. Taksonomičeskaja grammatika* [A structural description of the Archi language. V. 2. Taxonomic grammar]. Moscow: Moscow State University.

KIBRIK, ALEKSANDR (ed.). 1996. *Godoberi*. Munchen: Lincom Europa.

KIBRIK, A. E.; K. I. KAZENIN; E. A. LJUTIKOVA; and S. G. TATEVOSOV (eds.) 2001. *Bagvalinskij jazyk. Grammatika. Teksty. Slovari* [Bagwalal: grammar, texts, dictionaries]. Moscow: IMLI RAN.

PAYNE, JOHN R. 1995. Inflecting postpositions in Indic and Kashmiri. *Double Case: Agreement by Suffixaunahme*, ed. by Frans Plank, 283–98. Oxford: Oxford University Press.

SALIMOV, X. S. 2010. *Gagatlinskij govor andijskogo jazyka* [The Gagatli dialect of the Andi language]. Maxačkala: IJaLI RAN.

SPENCER, ANDREW. 2005. Case in Hindi. *Proceedings of LFG '05. CSLI Publications*, ed. by Miriam Butt, and Tracy H. King, 429–46. Online: http://csli-publications.stanford.edu/LFG/10/lfg05.html

SPENCER, ANDREW, and ANA R. LUÍS. 2012a. The canonical clitic. *Canonical morphology and syntax*, ed. by Dunstan P. Brown, Marina Chumakina, and Greville C. Corbett, 123–150. Oxford: Oxford University Press.

SPENCER, ANDREW, and ANA R. LUÍS. 2012b. *Clitics: An introduction*. Cambridge: Cambridge University Press.

SULEJMANOV, JA. G. 1957. *Grammatičeskij očerk andijskogo jazyka (po dannym govora s. Rikvani)* [A grammatical sketch of Andi based on the variant spoken in Rikvani]. Kand. diss., Institut jazykoznanija AN SSSR.

THAKUR, ANIL KUMAR. 1997. *Genitive construction in Hindi*. Master of philosophy thesis in linguistics. University of Delhi.

ZWICKY, ARNOLD M., and GEOFFREY K. PULLUM. 1983. Cliticization vs. inflection: English *n't*. *Language* 59.502–13.

Response Bias and Sensitivity in Distributional Learning: Two Stages in Early Phonetic Category Acquisition

EMILY MOENG
University of North Carolina at Chapel Hill[*]

1 Introduction

Despite being exposed to great variation and little to no explicit instruction, infants acquire linguistic structures with remarkable ease. Heart rate studies show that this acquisition begins even while infants are still within the womb (Minai et al. 2017). Newborns show recognition of their native language over other languages and recognition of their native language's prosodic structure over other prosodic structures (Mehler et al. 1988), as well as a preference for human speech over primate vocalizations or human speech played in reverse (Dehaene-Lambertz et al. 2002). By the time an infant becomes a year old, they exhibit language-specific discrimination of sounds which are contrastive in their language (Kuhl et al. 2006).

This study seeks to provide a **timeline of early segmental acquisition**. Specifically, concepts from Signal Detection Theory, RESPONSE BIAS and SENSITIVITY, will be related to distributional learning in a series of three experiments. It will be argued that phonetic category acquisition occurs in two stages: a Bias Stage followed by a Sensitivity stage. This runs contrary to models which base distributional learning in perceptual warping (Guenther & Gjaja 1996, Boersma et al. 2003), which do not predict a Bias Stage of phonetic category acquisition.

2 Background

This section begins by providing background on distributional learning in Section 2.1. Following this, Section 2.2 presents a distinction between response bias and sensitivity.

[*] This material is based upon work supported by the National Science Foundation under Grant No. 1451792. Special thanks to the P-side Research Group at the University of North Carolina, as well as the audiences at BLS 44 for their helpful comments and suggestions. All errors are my own.

2.1 Distributional learning and Maye and Gerken (2000)

Infants show language-specific discrimination of vowels around 6-8 months (Polka & Werker 1994), and language-specific discrimination of consonants at 6-12 months (Eilers et al. 1979). Although adult Japanese speakers experience considerable difficulty distinguishing between [ɹ] and [l] (Iverson et al. 2003), 6 month-old Japanese infants can discriminate between these two sounds (Kuhl et al. 2006). Similarly, although adult English speakers experience difficulty distinguishing between [t] and [ʈ], 8 month old English-learning infants are still able to distinguish these sounds (Werker & Tees 1984). Observations such as these lead to the claim that infants initially begin as 'citizens of the world' (Gervain & Mehler 2010, Kuhl 2004) with the ability to distinguish all contrasts which are linguistically-relevant, and that 'acquisition' essentially equates to a loss of contrasts which are not linguistically relevant in the language being heard (Gervain & Mehler 2010).[1] To borrow an analogy from Cristía et al. (2011), infant perception begins as a topographical map, with natural peaks separating speech sounds. Through experience, this initial topography is warped such that peaks are formed or existing peaks are flattened. The end result should be a language-specific perceptual map which aids in the discrimination of all contrastive phonemes in the target language.

How do infants acquire this language-specific map? One of the most widely-cited accounts for how this topography is warped is that language learners make use of distributional learning (Maye et al. 2000, Maye & Gerken 2002, Werker et al. 2012). According to this account, language learners map tokens into some phonetic space and make use of the frequency clusters that tokens make in this space to infer the number of phonetic categories in the language he or she is being exposed to (Maye & Gerken 2002). Learners exposed to a bimodal distribution of tokens along some phonetic dimension(s) will infer that there are two phonetic categories, whereas learners exposed to a monomodal distribution will infer that there is only one.

Maye and Gerken (2000) is the first artificial language learning study which provides experimental support for distributional learning. Maye and Gerken trained adult participants on syllables during an approximately 9-minute training phase. Critical syllables during this training phase began with an alveolar stop taken from an 8-point continuum ranging between a prevoiced

[1] For some exceptions to this seemingly universal discrimination shown by infants, see Polka et al. (2001), Sundara et al. (2006), and Narayan et al. (2010).

stop [d] and a voiceless unaspirated stop [d̥]. Participants were either trained on a bimodal distribution or a monomodal distribution of critical syllables (see Figure 1). For simplicity, continuum points will be referred to as D_1-D_8.

Following training, participants were directed to a test phase in which participants were presented with pairs of syllables and were asked if they believed the two syllables presented were the **same** word repeated twice, or two **different** words, in the language they had been exposed to. Maye and Gerken found a greater percentage of 'different' responses in the Bimodal condition than in the Monomodal condition when participants were presented with the endpoints of the continuum, D_1 and D_8. They conclude that this study supports distributional learning.

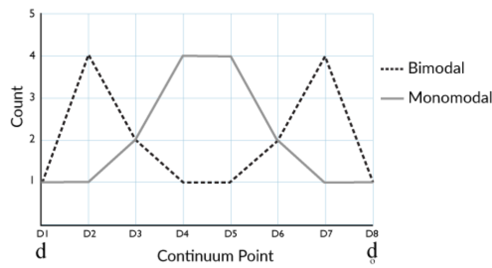


Figure 1. Illustration of the familiarization frequency of onsets of critical stimuli for Bimodal and Monomodal groups during the training phase of Maye and Gerken (2000).

Maye and colleague's findings are widely cited (for example, Kuhl 2004, Kuhl et al. 2006, Werker et al. 2012), and have been replicated a number of times. Experimental support has been found for adults (Maye & Gerken 2000, Maye & Gerken 2001, Hayes-Harb 2007) and infants (by measuring looking times) (Maye et al. 2002, Yoshida 2010).

2.2 Response bias vs. sensitivity

This section argues for a distinction to be made between two phenomena: RESPONSE BIAS and SENSITIVITY. To illustrate the difference between bias and sensitivity, suppose a subject is given the task of determining whether two marbles drawn randomly from a bag are the same color or are two different colors. Half of the participants are told that there exist two shades of green in the bag while the other half of the participants are not told anything about the number of shades of green. In reality though, there is only one shade of green marble. In this scenario, one might imagine that the participants who are told that there are two different shades of green in the bag are more likely to respond that two green marbles are different shades compared to participants who are not told anything about the number of shades of green in the bag, even though all participants are seeing the same exact same shade of green. In this first scenario, we can say that the group of participants told that there are two shades of green is more **biased** towards a 'different' response compared to the group of participants who had not been told anything regarding the number of shades of green in the bag. Even when faced with two marbles with the exact same shade, participants with a bias towards a 'different' response are more likely to respond 'different' compared to the participants who are not told anything about the number of shades of green, simply because they are expecting there to be two shades of green in the bag. Although this is a hypothetical example, factors which affect response bias have been identified (Macmillan & Creelman 2004).

We can imagine a second scenario, identical to the one above, except this time two different shades of green actually exist in the bag. In this scenario, one might imagine that the group that had been told that two different shades of green exist is actually better at detecting whether two green marbles are the same or different shades compared to the other group. That is, their SENSITIVITY to the slight difference in shade is actually heightened, because they know a difference exists, and their heightened awareness makes them more sensitive of the different shades. In this scenario, the group with the higher sensitivity responds 'different' more often when the two marbles are actually different shades of green, but, crucially also responds 'same' more often when the two marbles that are the same shade of green. In the previous scenario, the group with the lower response bias responded 'different' more often even when the two marbles were the same shade green. Sensitivity differences have been found in numerous tasks (e.g. see Sowden et al. 2000, Iverson & Kuhl 1995). A number of distributional learning studies find that groups trained on a bimodal distribution have a higher sensitivity (measured in d-prime) to critical stimuli after training than a monomodal or control group (Hayes-Harb 2007, Noguchi 2016).

The scenario of drawing two marbles from a bag is similar to that which is presented to a participant in a distributional learning experiment. During the test phase, the participant is asked to determine whether two sounds, rather than marbles, are the same or different. One can imagine two different theories regarding the underlying mechanism behind distributional learning, each of which make different predictions regarding sensitivity and response bias, stated below.

1) **The Sensitivity Hypothesis**: Distributional learning affects sensitivity, such that a bimodally-trained learner perceives phones more distinctly than a monomodally-trained learner (see Figure 2).

2) **The Response Bias Hypothesis**: Bimodal training results in learners being more likely to think that two sound categories exist in the speech stream compared to monomodal training. Therefore, the bimodal learners have a bias towards responding 'different.' Perception, however, is not directly affected by distributional learning (see Figure 2).

When sensitivity increases, it is expected that learners will be better at determining that two different stimuli are in fact different, as well as better at determining that two identical stimuli are the same. However, when a bias for responding 'different' increases, it is expected that learners determine that pairs of stimuli are different more often, regardless of whether they are in fact the same or different. This is summarized in Table 1.

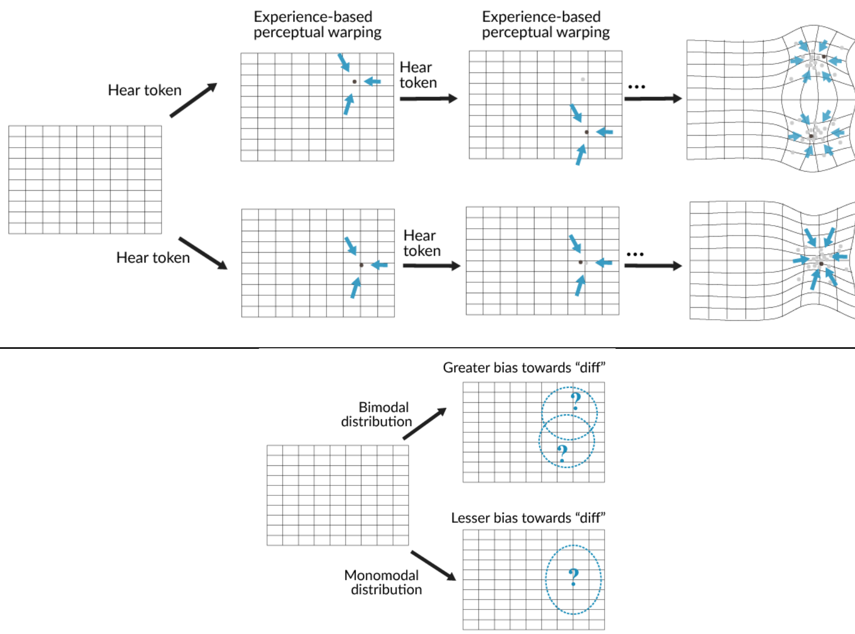


Figure 2. Illustration of the Sensitivity Hypothesis (top) and Response Bias Hypothesis (bottom).

Previous metrics used to measure distributional learning either measure sensitivity or are inconclusive in whether they measure sensitivity or response bias. Maye and Gerken (2000) analyze the percentage of times participants respond that critical Different Pairs are 'different', finding that the participants trained on a monomodal distribution respond 'different' less often than participants trained on a bimodal distribution. However, as can be seen in Table 1, this finding is compatible with an increase in sensitivity *or* in a decrease in response bias. Since Maye and Gerken only analyze the effect on Different Pairs, it is unclear whether sensitivity or response bias is the factor being affected. Noguchi (2016) measures sensitivity in the form of d-prime values. He finds that participants in the control condition have lower d-prime values than participants trained on a bi-modal distribution. Since d-primes also take responses for Same Pairs into account, this study is able to distinguish between sensitivity and response bias.

The distinction between sensitivity and response bias is an important one to make, because perceptual warping accounts of distributional learning (such as those suggested by Guenther and Gjaja (1996) and Boersma et al. (2003)) predict that sensitivity will **always** be affected by distributional learning, since a change in sensitivity is the underlying cause of distributional learning in these models (see Moeng (2018) for details). The current study measures both sensitivity and bias, and concludes that both can be affected by distributional learning. I argue that direct perceptual warping accounts of distributional learning are not supported by this study, and that a change in response bias occur in an initial stage of distributional learning, which can be followed by a change sensitivity occurs at a second stage.

	Percentage of 'different' responses for Different Pairs (e.g. D_1a vs. D_8a)	Percentage of 'different' responses for Same Pairs (e.g. D_1a vs. D_1a)
Increased bias towards 'different' response	↑	↑
Increased sensitivity	↑	↓

Table 1. Effect of increased bias towards a 'different' response compared to effect of increased sensitivity on a learner's responses.[2]

3 Experiment 1[3]

3.1 Stimuli

Stimuli consisted of critical syllables and filler syllables. Onsets of critical syllables were drawn from three 8-point continua ranging between voiceless unaspirated [k] (*s**k̲**ill*), and prevoiced [g] (***g̲**ill*). Continuum points will be referred to as G_1-G_8, where G_1 indicates the most [g]-like end of the continuum, and G_8 indicates the most [g̊]-like end. Following Maye and Gerken (2000), each of the three continua differed in following nucleus: [g̊ɑ]-[gɑ], [g̊æ]-[gæ], and [g̊ɹ̩]-[gɹ̩].

---

[2] The third scenario, in which both response bias and sensitivity are affected, is not discussed here.

[3] Please see Moeng (2018) for details regarding the procedure of Experiments 1-3, including stimuli creation and participant recruitment/inclusion criteria.

3.2 Participants

Participants were recruited from Amazon's Mechanical Turk ('MTurk'). They were asked to participate only if they had no known history of speech or hearing impairments and were a native speaker of English. Because this experiment was conducted online rather than face-to-face, only participants using a computer in the United States were allowed to participate to increase the chance that the participant would be a native English speaker.[4]

3.3 Procedure

Experiment 1 consisted of a Practice Test, a Train phase, a Test phase, and a Questionnaire[5]. In the Practice Test, participants were then presented with pairs of English words produced by the same speaker that were either Same Pairs, or Different Pairs. Same Pairs consisted of repetitions of the same word that were different enough to be distinguished as different tokens (e.g. $lock_1$ vs. $lock_2$). Different Pairs consisted of minimal pairs (e.g. *lock* vs. *rock*, *desk* vs. *disk*). Participants were asked to press the 'S' key if the pairs of words that they heard were the 'same' word, or the 'D' key of they were 'different' words. Responses were scored as 'correct' if participants responded 'different' on Different Pairs and 'same' on Same Pairs. Participants who answered fewer than 5/8 correct on the Practice Test were excluded. No participants failed to meet this.

During the Train phase, participants heard a monomodal or bimodal distribution of phones, depending on which condition they were in. The Bimodal group heard a bimodal frequency of phones; the Monomodal group heard a monomodal frequency distribution of phones (see Figure 10. In addition, three recordings of 8 filler syllables ([fɑ], [fæ], [tɛ], [tej], [mæ], [næ], [sɛ], and [zɛ]) were made. Each filler was repeated twice during each Train repetition. Each repetition was repeated 4 times, resulting in 192 fillers and 192 critical tokens.

During the Test phase, participants were given pairs of syllables that were either Same Pairs, or Different Pairs. Same Pairs consisted of repetitions of the same exact token for critical tokens (i.e. G_1a vs. G_1a), or different tokens for control tokens ($[\text{fɑ}]_1$ vs. $[\text{fɑ}]_2$). Control Same Pairs were judged by the experimenter to sound different enough to be distinguished as separate tokens. Critical Different Pairs consisted of pairs that occurred on opposite ends of the 8-point continuum

[4] This can be done through MTurk 'qualifications.' See Moeng (2018) for other qualifications used.
[5] This was preceded by a Sound Check to determine whether participants were wearing headphones and paying attention. See Moeng (2018) for details.

(i.e. G_1a vs. G_8a). Critical Same Pairs consisted of identical tokens on the continuum (i.e. G_1a vs. G_1a, G_8a vs. G_8a). Participants were asked to press the 'D' key if they believed the two syllables were different words and the 'S' key if they believed they were repetitions of the same word. Each Test repetition consisted of 4 critical pairs (2 Same Pairs and 2 Different Pairs), for each vowel context, resulting in 12 critical pairs. Each Test repetition also contained 4 control Same Pairs and 4 control Different Pairs. There were two repetitions of each Test phase resulting in a total of 24 critical pairs and 16 control pairs. In total, 27 participants in the Bimodal group and 34 in the Monomodal group were included in the analysis.

3.4 Analysis

The regression formula described below models one dependent variable, RESPONSE, with two fixed effects: (1) DISTRIBUTION, a between-subject, within-item factor consisting of two levels {*bi-modal*, *monomodal*}and (2) PAIRTYPE, a within-subject, between-item factor consisting of two levels {*same*, *diff*}. This was done separately for *critical* and *control* trials. The dependent variable RESPONSE consisted of two levels, *s* and *d*, where *s* corresponds to a participant response of 'same' during the Test phase, and where *d* corresponds to a participant response of 'different' during the Test phase. Random effects for SUBJECT and ITEM were also included in the model.

All statistical tests were completed in R (R Core Team, 2015), using the `glmer` function from the `lme4` package (Bates et al. 2015) to fit a generalized linear mixed-effects model (GLMM) with a logistic link function ('mixed logit model'). Significance was set at a level of $p < 0.05$. Two separate regressions were conducted to compare the effect of Distribution and PairType on Response: one for *critical* items, and one for *control* items. The following formula was used:

(1) Response ~ Distribution*PairType + (1+PairType|Subject) + (1+Distribution|Item)

The results of the mixed logit model are interpreted in the following way: a significant main effect of Distribution without a significant interaction between Distribution and PairType is interpreted as evidence for a difference in **bias** between the two groups of participants. That is, if participants trained on Distribution X have greater bias to respond 'different' than participants trained on Distribution Y, we would expect to see participants trained on Distribution X respond 'different' more often than participants trained on Distribution Y for all trials, regardless of whether or not the trial consisted of a Same Pair or a Different Pair. We expect to see this in the results of a fitted model as a significant main effect of Distribution, but **not** as a significant interaction between Distribution and PairType since, in this scenario, the effect of PairType would not differ depending on which Distribution participants were exposed to. An interaction of Distribution and PairType will be interpreted as a difference between groups of participants in sensitivity to the slight distinction between G_1 and G_8. That is, if participants trained on Distribution X are more sensitive to the acoustic differences between G_1 and G_8 compared to participants trained on Distribution Y, we would expect participants trained on Distribution X to respond 'different' more often than participants trained on Distribution Y for Different Critical trials, but to also respond 'different' **less** often for Same Critical trials, resulting in a significant interaction between Distribution and PairType.

3.5 Results

Summaries of the fixed effects in the mixed logit model with treatment coding are shown in the Appendix. No interaction between Distribution and PairType was found for critical trials (p = 0.650), but an interaction was found for control trials (p = 0.047). To find the main effect of Distribution on Response, a planned contrast analysis was performed on the fitted GLMM using the `glht` function from the `multcomp` package (Hothorn et al. 2017) in R. The planned contrast revealed a main effect of Distribution for critical trials (p = 0.049), with the Bimodal group having greater odds of responding *d* than the Monomodal group. No main effect was found for control trials (p = 0.817).

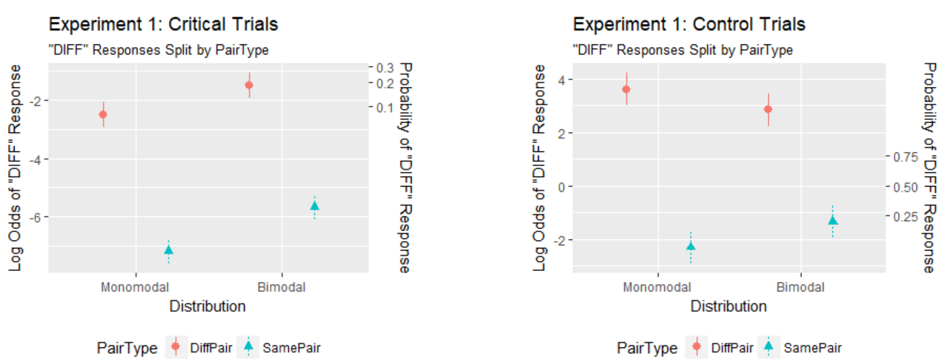


Figure 3. Log odds of participants responding 'different', Experiment 1.

4 Experiment 2

The goal of Experiment 2 is to determine whether the above results are particular to the stimuli used in Experiment 1. Experiment 2 follows the methodology of Maye and Gerken (2001) and

Hayes-Harb (2007) as closely as possible, using stimuli obtained from Maye and Gerken. The procedure also closely follows that used in Maye and Gerken (2001), with the small addition of the Sound Check task preceding the experiment.[6] For reasons of space, details will not be provided here, but can be found in Moeng (2018, Chapter 3). We again find similar results as in Experiment 1: we find no interaction between Distribution and PairType for Critical trials (p = 0.114), or for Control trials (p = 0.228), but we do find a main effect of Distribution for critical trials (p = 0.033), with the Bimodal group having greater odds of responding *d* than the Monomodal group. No main effect was found for control trials (p = 0.934).

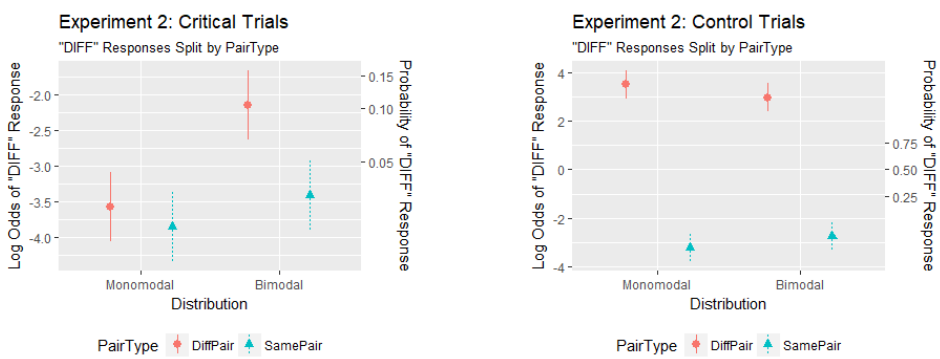


Figure 4. Log odds of participants responding ‘different’, Experiment 2.

5 Experiment 3

The previous experiments found evidence that bimodally-trained participants have a greater bias towards a ‘different’ response compared to monomodally-trained participants. The goal of Experiment 3 is to determine whether we find evidence that bimodal training causes an increase in **sensitivity** compared to monomodal training if we use critical stimuli which range between more perceptually distinct endpoints. This experiment followed the general procedure of Experiments 1 and 2, but used the critical stimuli [ɕɑ] – [ʂɑ]. Experiment 3 excluded participants who had experience with more than one voiceless post-alveolar fricatives as phonemes, following Noguchi (2016) who also used this contrast as critical stimuli. For participant details, see Moeng (2018). In total, the number of participants included in the analysis was 22 participants in the Bimodal group and 27 in the Monomodal group.

[6] Many thanks to Rachel Hayes-Harb, LouAnn Gerken, and Jessica Maye for sending me and allowing me to use their stimuli.

5.1 Stimuli

Stimuli consisted of critical syllables and filler syllables. Following Noguchi (2016), onsets of critical syllables were drawn from an 8-point continuum ranging between an alveopalatal fricative [ɕ] to a retroflex fricative [ʂ]. Continuum points will be referred to as S_1-S_8, where S_1 indicates the most [ɕ]-like end of the continuum, and S_8 indicates the most [ʂ]-like end.[7]

5.2 Results

A GLMM was fitted to the formula in (1). For Experiment 3, an interaction between Distribution and PairType was found for Critical trials ($p = 0.037$). Surprisingly, an interaction between Distribution and PairType was found for Control trials as well ($p = 0.009$). See Appendix for details.

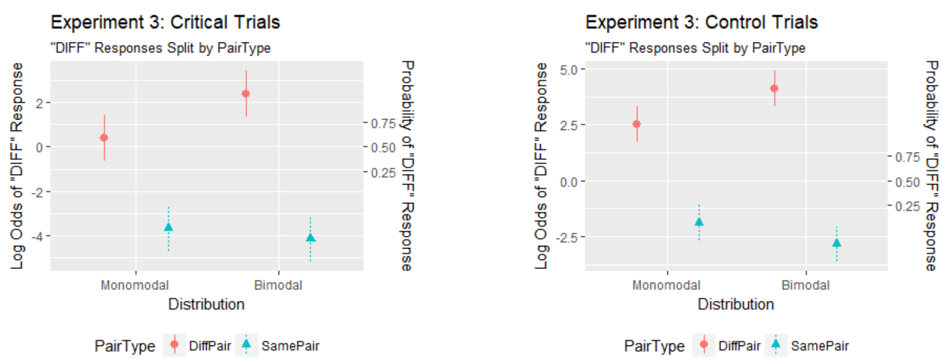


Figure 5. Log odds of participants responding ‘different’, Experiment 3.

5.3 Discussion

Despite the significant interaction unexpectedly found between Distribution and PairType for control trials, this study concludes that stimuli used in Experiment 3 were able to induce greater sensitivity in a bimodally-trained group of participants compared to a monomodally-trained group of participants, since a significant interaction between Distribution and PairType was found for critical trials. It is unknown why the control trials also exhibited a significant interaction.[8]

[7] Although previous experiments followed critical onsets with three different rimes ([ɑ æ ɚ]), the current experiment follows Noguchi (2016) and all onsets are only followed by [ɑ].

[8] Brief speculation is provided in Moeng (2018).

6 Discussion

This study finds evidence for distributional learning through an online platform, in bias and in sensitivity. Experiments 1 and 2 find that the bimodal group has a greater bias to respond 'different' than the monomodal group, while Experiment 3 finds that the bimodal group has a higher sensitivity than the monomodal group.

6.1 Proposal

In Section 2.2, two hypotheses regarding the driving mechanism behind distributional learning were presented: the Sensitivity and Bias Hypotheses. Models which base distributional learning in perceptual warping (Guenther & Gjaja 1996, Boersma et al. 2003), assume a sensitivity-based account. The set of experiments conducted in this study finds support for sensitivity changes, as well as changes in bias.

This study proposes a two-stage model for the acquisition of phonetic categories. In the first stage, learners form a rough idea of the number of phonetic categories in the speech stream. During this stage, the distribution of phones encountered affects learners' biases, such that exposure to a bimodal distribution of phones causes learners to expect two phonetic categories and a monomodal distribution of phones causes learners to expect one phonetic category (see Figure 6).

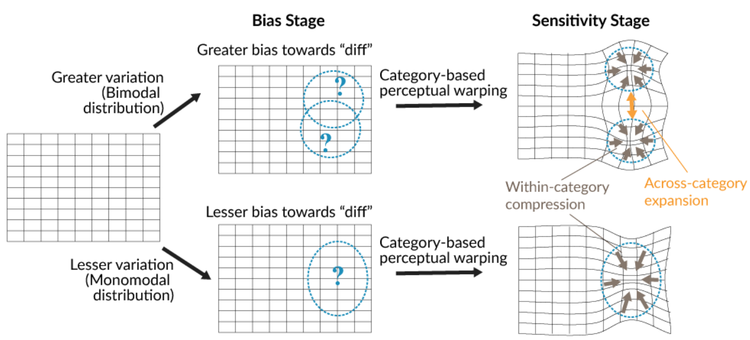


Figure 6. Illustration of proposed mechanism behind phonetic category acquisition.

In this Bias Stage, the learner only has a general idea of the number of phonetic categories, and does not have well-defined category centers or category boundaries.

What is responsible for the difference in bias between the bimodally- and monomodally-trained language learners? Although further research is required to determine this, we believe this can be explained by how much participants notice variation in the training phase. Since the bimodal group hears more tokens at the continuum points 2 and 7 and the monomodal group hears more tokens at continuum points 4 and 5, it is more likely that a participant during bimodal training will encounter continuum point 2 followed by continuum point 7 (or vice versa) than a participant during monomodal training, who is more likely to encounter continuum point 4 followed by continuum point 5. Because continuum points 2 and 7 are more acoustically distinct, a bimodally-trained participant may be more awareness to the difference between critical stimuli than a monomodally-trained participant, simply because they hear two distinct tokens presented side-by-side more often than the monomodal participants. Even if the bimodal participants are not able to identify what the difference between these two sounds might be, we could imagine they might have a stronger awareness that there at least **are** two different sounds compared to the monomodal participants, resulting in more ‘different’ responses when the bimodal participants are presented with pairs of critical syllables. One interesting way to test this would be to train participants on a bimodal or monomodal distribution along one phonetic dimension, but then test participants with pairs of sounds which differ along a different phonetic dimension. If participants with a greater ‘different’ bias believe there are two different sounds but truly do not know how they differ, we would expect these participants to respond that pairs of sounds are different even along an untrained dimension.[9] Alternatively, we could test this by manipulating the order in which a bimodal distribution is presented to participants, either in a way that maximizes ‘2-7’ back-to-back pairs (e.g. ‘2… 7… 2…’), or in a way that minimizes them (e.g. ‘1… 2… 2… 2… 3…’).

Supposing we are able to explain differences in bias as an awareness of variation in the speech signal, we would still need to explain how some distributional learning experiments result in greater sensitivity in bimodally-trained participants compared to monomodally-trained participants. We could potentially explain this second stage with some perceptual warping distributional learning mechanism such as those proposed by Guenther and Gjaja (1996) or Boersma et al. (2003), but we believe this can be explained with domain-general mechanisms which have already been thoroughly documented in the psychology literature: **across-category expansion** and **within-category compression** (Livingston et al. 1998, Goldstone 1994). That is, the formation of some category warps perception such that items within a category are perceived as being more similar

[9] Many thanks to Stephan Gries for this interesting suggestion for future research.

to one another, and items from different categories are perceived as being more dissimilar from one another.

To summarize, we propose that learners exposed to a bimodal distribution of phones are presented with more examples of noticeably-different stimuli. This leads to an increase in awareness that there are two different sounds in a BIAS STAGE of distributional learning. Learners who are more aware that there are two different sounds are more likely to attempt to categorize these sounds into different categories, and category-based perceptual warping results in greater sensitivity to sounds which cross category boundaries in a SENSITIVITY STAGE. Neither of these stages require the proposal of a special mechanism for phonetic category acquisition.

7 Conclusion

In a set of three artificial language learning experiments conducted through Mechanical Turk[10], this chapter makes a distinction between bias and sensitivity, and proposes a model in which phonetic category acquisition occurs in two stages: first, learners form rough bias regarding the number of sound categories in the speech stream in a Bias Stage. Following this, learners' sensitivities change through general cognitive mechanisms which drive within-category compression and across-category expansion (Goldstone 1994) in a Sensitivity Stage. This two-stage model of phonetic category acquisition counters predictions made by previous accounts of distributional learning (Guenther & Gjaja 1996, Boersma et al. 2003), which predict that a change in sensitivity always accompanies distributional learning.

REFERENCES

BATES, DOUGLAS, MARTIN MAECHLER, BEN BOLKER, and STEVEN WALKER. 2015. "lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014."

BOERSMA, PAUL, PAOLA ESCUDERO ESCUDERO, and RACHEL HAYES. 2003. "Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories." *15th International Congress of Phonetic Sciences.* 1013-1016.

CRISTIÀ, ALEJANDRINA, GRANT L MCGUIRE, AMANDA SEIDL, and ALEXANDER L FRANCIS. 2011. "Effects of the distribution of acoustic cues on infants' perception of sibilants." *Journal of Phonetics* 39 (3): 388-402.

[10] Results here are based on experiments performed on adults. See Chapter 3 of Moeng (2018) for evidence based on reported non-replications of Maye and Gerken (2002) in the past infant study of Yoshida et al. (2010).

DEHAENE-LAMBERTZ, GHISLAINE, STANISLAS DEHAENE, and LUCIE HERTZ-PANNIER. 2002. "Functional neuroimaging of speech perception in infants." *Science* 298 (5600): 2013-2015.

GERVAIN, JUDIT, and JACQUES MEHLER. 2010. "Speech perception and language acquisition in the first year of life." *Annual Review of Psychology* 61: 191-218.

GOLDSTONE, ROBERT L, and ANDREW T HENDRICKSON. 2010. "Categorical perception." *Wiley Interdisciplinary Reviews: Cognitive Science* 1 (1): 69-78.

GUENTHER, FRANK H, and MARIN N GJAJA. 1996. "The perceptual magnet effect as an emergent property of neural map formation." *The Journal of the Acoustical Society of America* 100 (2): 1111-1121.

HAYES-HARB, RACHEL. 2007. "Lexical and statistical evidence in the acquisition of second language phonemes." *Second Language Research* 23 (1): 65-94.

IVERSON, PAUL, and PATRICIA K KUHL. 1995. "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling." *The Journal of the Acoustical Society of America* 97 (1): 553-562.

IVERSON, PAUL, PATRICIA K KUHL, REIKO AKAHANE-YAMADA, EUGEN DIESCH, YOH'ICH TOHKURA, ANDREAS KETTERMANN, and CLAUDIA SIEBERT. 2003. "A perceptual interference account of acquisition difficulties for non-native phonemes." *Cognition* 87 (1).

KUHL, PATRICIA K. 2004. "Early language acquisition: cracking the speech code." *Nature Reviews Neuroscience* 5 (11): 831-843.

KUHL, PATRICIA K, ERICA STEVENS, AKIKO HAYASHI, TOSHISADA DEGUCHI, SHIGERU KIRITANI, and PAUL IVERSON. 2006. "Infants show a facilitation effect for native language phonetic perception between 6 and 12 months." *Developmental Science* 9 (2).

LIVINGSTON, KENNETH R, JANET K ANDREWS, and STEVAN HARNAD. 1998. "Categorical perception effects induced by category learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24 (3): 732.

MACMILLAN, NEIL A, and C. DOUGLAS CREELMAN. 2004. *Detection theory: A user's guide.* 2nd. Mahwah, NJ: Lawrence Erlbaum Associates.

MAYE, JESSICA, and LOUANN GERKEN. 2000. "Learning phonemes without minimal pairs." *24th Annual Boston University Conference on Language Development.*

Maye, Jessica, LouAnn Gerkekn. 2001. "Learning phonemes: How far can the input take us." *25th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press.

MAYE, JESSICA, JANET F WERKER, and LOUANN GERKEN. 2002. "Infant sensitivity to distributional information can affect phonetic discrimination." *Cognition* 82 (3): B101-B111.

MEHLER, JACQUES, PETER JUSCZYK, GHISLAINE LAMBERTZ, NILOFAR HALSTED, JOSIANE BERTONCINI, and CLAUDINE AMIEL-TISON. 1988. "A precursor of language acquisition in young infants." *Cognition* 29 (2): 143-178.

MINAI, UTAKO, KATHLEEN GUSTAFSON, ROBERT FIORENTINO, ALLARD JONGMAN, and JOAN SERENO. 2017. "Fetal rhythm-based language discrimination: a biomagnetometry study." *NeuroReport* 28 (10): 561-564.

MOENG, EMILY. 2018. *The Acquisition of Phonetic Categories.* Chapel Hill, NC: The University of North Carolina at Chapel Hill.

NOGUCHI, MASAKI. 2016. *Acquisition of allophony from speech input by adult learners.* Vancouver: The University of British Columbia.

POLKA, LINDA, and JANET F WERKER. 1994. "Developmental changes in perception of nonnative vowel contrasts." *Journal of Experimental Psychology: Human Perception and Performance* 20 (2): 421.

SOWDEN, PAUL T, IAN R. L DAVIES, and PENNY ROLING. 2000. "Perceptual learning of the detection of features in X-ray images: A functional role for improvement in adults' visual sensitivity?" *Journal of Experimental Psychology: Human Perception and Performance* 26 (1): 379-390.

WERKER, JANET F, and RICHARD C TEES. 1984. "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life." *Infant Behavior and Development* 7 (1): 49-63.

WERKER, JANET F, H. HENNY YEUNG, and KATHERINE A YOSHIDA. 2012. "How do infants become experts at native-speech perception?" *Current Directions in Psychological Science* 21 (4): 221-226.

YOSHIDA, KATHERINE A, FERRAN PONS, JESSICA MAYE, and JANET F WERKER. 2010. "Distributional phonetic learning at 10 months of age." *Infancy* 15 (4): 420-433.

Appendix

Predictor	Coefficient	SE	Wald Z	p
Critical				
(Intercept)	-1.502	0.405	-3.707	<0.001 ***
Distribution=*monomodal*	-0.990	0.434	-2.280	0.023 *
PairType=*same*	-4.174	1.652	-2.526	0.012 *
Interaction=monomodal & same	-0.513	1.130	-0.454	0.650
Control				
(Intercept)	2.834	0.517	5.486	<0.001 ***
Distribution=*monomodal*	0.769	0.605	1.271	0.204
PairType=*same*	-4.169	0.709	-5.881	<0.001 ***
Interaction=monomodal & same	-1.743	0.877	-1.988	0.047 *

Table 2. Summary of fixed effects in the mixed logit model in Experiment 1.

Predictor	Coefficient	SE	Wald Z	p
Critical				
(Intercept)	-2.152	0.318	-6.773	<0.001 ***
Distribution=*monomodal*	-1.420	0.485	-2.927	0.003 **
PairType=*same*	-1.255	0.471	-2.667	0.008 **
Interaction=monomodal & same	0.982	0.622	1.579	0.114
Control				
(Intercept)	2.977	0.471	6.316	<0.001 ***
Distribution=*monomodal*	0.529	0.568	0.932	0.351
PairType=*same*	-5.720	0.673	-8.503	<0.001 ***
Interaction=monomodal & same	-1.001	0.830	-1.206	0.228

Table 3. Summary of fixed effects in the mixed logit model in Experiment 2.

Predictor	Coefficient	SE	Wald Z	p
Critical Trials				
(Intercept)	2.365	0.809	2.922	0.003 **
Distribution=*monomodal*	-1.978	1.040	-1.903	0.057
PairType=*same*	-6.489	0.963	-6.738	<0.001 ***
Interaction=monomodal & same	2.439	1.166	2.091	0.037 *
Control Trials				
(Intercept)	4.112	0.759	5.415	<0.001 ***
Distribution=*monomodal*	-1.599	0.780	-2.052	0.040 *
PairType=*same*	-6.957	0.944	-7.371	<0.001 ***
Interaction=monomodal & same	2.549	0.978	2.606	0.009 **

Table 4. Summary of fixed effects in the mixed logit model in Experiment 3.

Interpretation of internally-headed relative clauses in Shan

MARY MORONEY
Cornell University[*]

1 Introduction

Shan is a Southwestern Tai language related to Thai that is spoken in parts of Myanmar, Thailand, and other countries (Glick & Moeng 1991). In this paper, I work through evidence showing the existence of internally headed relative clauses (IHRCs) in Yǔn Shan, a dialect of Shan, and highlight an indefinite interpretation of IHRCs in this language that does not fit well with previous IHRC typologies. The basic IHRC and Post-Head constructions as shown in 1.[1] In 1a is an IHRC. The head of the clause, *lik* 'book', appears in the same place it would in independent clause SVO structure. In 1b, on the other hand, the head is in the typical Post-Head position, appearing before the whole clause, in front of *ʔǎn*, which I analyze as a complementizer.

(1) **Relative clauses in Shan**

a. [ʔǎn háw hǎn **lik** nâj] mán lɛ̌ŋ.
COMP 1.SG see book this 3.SG red
'The/A book that I see is red.' **(IHRC)**

b. [**lik** [ʔǎn háw hǎn] nâj] mán lɛ̌ŋ.
book COMP 1.SG see this 3.SG red
'The/A book that I see is red.' **(Post-Head)**

*Thanks to Aye Twei Soe who provided the Shan data. Thanks also to Molly Diesing, Carol-Rose Little, Sarah Murray, and John Whitman and the audiences of the Chulalongkorn International Student Symposium on Southeast Asian Linguistics and BLS 44 for all their feedback. Any errors are my own.

[1]The Shan data comes from my fieldwork with a Shan speaker in Ithaca, NY from January 2016 to September 2017. My consultant is from Meiwai village, near Papun in Kayin (Karen) State in Myanmar. She speaks the Yǔn Shan dialect, which is very different from the Taunggyi dialect. She also speaks Karen, Burmese, and English. She has been in the United States for 5 years. Data was collected using a variety of elicitation methods: telling short stories, grammaticality judgments, and felicity judgments.

Glossing conventions: 1: first person, 3: third person, ACC: accusative, ANIM: animal, ASP: aspect, CL: classifier, COMP: complementizer, GEN: genitive, HUM: human, IMPF: imperfect, IRR: irrealis, LOC: locative, NOM: nominative, PAST: past tense, PERF: perfect, PL: plural, ROUND: round, SG: singular, SM: scope marker, TOP: topic, VAL: validator

The availability of IHRCs in Shan is somewhat typologically unusual because SVO languages are less likely to have internally headed relative clauses: Cole (1987) had proposed that IHRCs were only found in OV languages. Recently, IHRCs have been identified in verb initial languages like Seediq and Tagalog (Aldridge 2004) and SVO like Buli (Hiraiwa 2003). WALS counts 58/580 OV languages and 5/608 VO languages as having IHRCs (Dryer & Haspelmath 2013).

Section 2 gives several examples of internally headed relative clauses in Shan. Section 3 goes over the previously described typology of internally headed relative clauses, focusing on the issues of maximality and island sensitivity. Section 4 demonstrates that Yǔn Shan IHRCs are island sensitive and non-maximal which complicates the typology of IHRCs. Section 5 describes the IHRC data of Navajo, which seems to have relative clauses of the same type as Shan. Section 6 sketches a head raising analysis and concludes.

2 Yǔn Shan relative clauses

Example 2 shows the strict SVO word order of Shan. *Lik* 'book' in 2 could be definite or indefinite because Shan lacks overt articles. As 3 shows, it is also a classifier language.

(2) háw hǎn lik.
1.SG see book
'I see the/a book.'

(3) háw hǎn mǎa sǎam tǒ.
1.SG see dog three CL.ANIM
'I see three dogs.'

In order to distinguish between internally and externally headed relatives, I use: (i) the complementizer, *ʔǎn*, to identify the left edge of the clause, and (ii) *wánnâj* 'today', to identify the right edge of the clause. In 4a, the head *kàj* 'chicken' appears in the usual position for objects. It is clearly inside of the clause because *wánnâj* 'today', which is also inside the relative clause, appears to the right of the head. In 4b, the head is clearly outside of the relative clause since it appears to the left of the complementizer *ʔǎn*.

(4) Extracted object as matrix subject

a. [ʔǎn Nan Li sɯ́ **kàj** wánnâj] mán pěn sǐ khǎaw.
COMP Nan Li bought chicken today 3.SG be color white
'The/A chicken Nan Li bought today was white.' (**IHRC**)

b. [**kàj**$_i$ [ʔăn Nan Li sɯ́ɯ t_i wánnâj]] mán pĕn sǐ khăaw.
chicken COMP Nan Li bought today 3.SG be color white
'The/A chicken Nan Li bought today was white.' (**Post-Head**)

Examples 4a-b show that the head of the relative clause can function as the matrix subject, and 5a and 5b show examples where the head is functioning as the matrix object. From these examples it is possible to see that the Shan IHRC is not simply a topicalization structure.

(5) **Extracted object as matrix object**

a. Saj Kham ʔăw [ʔăn Nan Li sɯ́ɯ máa **kàj** wánnâj].
Saj Kham take COMP Ms. Li buy come chicken today
'Saj Kham took the/a chicken Nan Li bought today.' (**IHRC**)

b. Saj Kham ʔăw [**kàj**$_i$ [ʔăn Nan Li sɯ́ɯ máa t_i wánnâj]].
Saj Kham take chicken COMP Ms. Li buy come today
'Saj Kham took the/a chicken Nan Li bought today.' (**Post-Head**)

Many of my examples were elicited and tested in a constructed context, but 6a is an example from a fable my consultant told. Here the head of the relative clause is *sɤkhó hàw* 'clothes', and it appears in typical object position rather than in front of the relative in a post-head construction.

(6) **Extracted object: Example from story**

a. Luk kɔ̂ kăaŋ nâjhankɔ mán sɯ́ɯ sàw [ʔăn mán cɤ̌k **sɤkhó hàw**].
child CL.HUM middle TOP he buy put COMP he like clothes PL
'The middle child, he bought and put the clothes that he liked,' (**IHRC**)

It is also possible to extract the subject of the relative clause, as in 7a-b, or the object of a ditransitive as in 8a-b.

(7) **Extracted Subject**

a. [ʔăn **kàj** cǐn khaw jù nâj] mán pĕn sǐ khăaw.
COMP chicken eat rice IMPF this 3.SG be color white
'The/A chicken eating rice is white.' (**IHRC**)

b. [**kàj**$_i$ [ʔăn t_i cǐn khaw jù] nâj] mán pěn sǐ khăaw.
chicken COMP eat rice IMPF this 3.SG be color white
'The/A chicken eating rice is white.' (**Post-Head**)

(8) **Extracted Object of a Ditransitive**

a. Saj Kham ʔăw [ʔăn Nan Li păn lukʔɔ̀n **kàj** wánnâj].
Saj Kham take COMP Nan Li give child chicken today
'Saj Kham took the/a chicken Nan Li gave to the child today.' (**IHRC**)

b. Saj Kham ʔăw [**kàj**$_i$ [ʔăn Nan Li păn lukʔɔ̀n t_i wánnâj]].
Saj Kham take chicken COMP Nan Li give child today
'The/A chicken eating rice is white.' (**Post-Head**)

As this section shows, this variety of Shan does have internally headed relative clauses, and it is possible to extract from a variety of argument positions.

3 Island sensitivity and maximality: IHRC typology

Some analyses of IHRCs have noted an apparent correlation between maximal interpretations and wh-island constraints (Grosu 2002, Watanabe 2004), leading them to categorize IHRCs as one of two types: (i) the Lakhota type: restrictive/non-maximalizing, island insensitive; and (ii) the Quechua and Japanese type: maximalizing, island sensitive.

3.1 Islands: CNPC

Lack of island sensitivity has typically been used as evidence that syntactic A'-movement is not taking place in these constructions and that instead binding is responsible for the available interpretation (Bonneau 1990, Grosu 2000, Watanabe 2004). Languages like Lakhota are not sensitive to islands, as 9 shows. Here, the head, *wowapi* 'paper', is being extracted out of a relative clause that has a different head (i.e. *wichota* 'many people'). Languages like Japanese have IHRCs that are sensitive to islands, as shown in 10. In this example, the head, *subarashii ronbun* 'excellent paper', is being extracted out of a relative clause whose head is *hito* 'person'. This extraction out of another relative clause is not possible in Japanese, as it violates the Complex Noun Phrase Constraint (CNPC) identified by Ross (1967).

(9) [[Wichota **wowapi** wą yawa pi cha] ob woʔųglaka pi ki] he L.A. Times e.
many-people paper a read PL ind with we-speak PL the that L.A. Times be
'The newspaper that we talk to many people who read (it) is the L.A. Times.'
(**Lakhota**; Williamson 1987, cited in Watanabe 2004: (8))

(10) *[John-ga [**subarashii ronbun**-o kaita hito]-o homete-ita no]-ga
John-NOM excellent paper-ACC wrote person-ACC praised-had COMP-NOM
shuppan-sareta.
publish-PASS
'An excellent paper which John had praised the person who wrote (it) was published.'
(**Japanese**; Watanabe 2004: (4b))

Since languages fall neatly into these patterns, researchers have explained the difference in sensitivity to the CNPC using the differences in on how the relative clause is formed: often, movmenet versus binding. This paper focuses on languages that display island sensitivity.

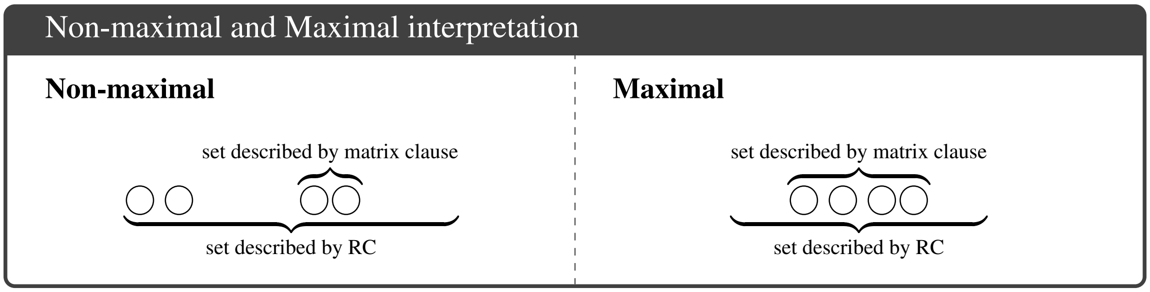
3.2 Maximality

An internally headed relative clause can be interpreted maximally or non-maximally. The distinction between a maximal and non-maximal interpretation corresponds approximately to having a definite determiner 'the' modifying the head of the relative in English, as shown in 11-12. 11 has a non-maximal interpretation because it is not the case that Bill must have eaten all of the apples Susan bought. 12 has a maximal interpretation: it would be infelicitous unless Bill ate all of the apples Susan bought.

(11) Bill ate apples that Susan bought.

(12) Bill ate the apples that Susan bought.

The figure below summarizes these interpretations. The maximal interpretation is obligatory when the set described by the relative clause (e.g., the apples that Susan bought) has to be co-extensive with the set described by the matrix clause (the apples eaten by Bill).The maximal interpretation is the case where all of the elements in the relative clause must be in the set described by the matrix clause.

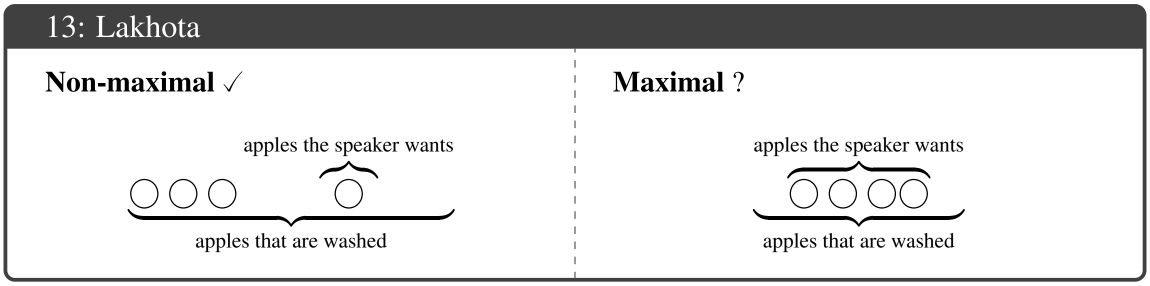


Analyses derive the maximal interpretation of IHRCs in various ways: an e-type analysis (Shimoyama 1999), feature checking with a null D of a wh-phrase (Watanabe 2004), quantificational disclosure (Grosu & Landman 2012), and trace conversion (Erlewine & Gould 2016). Grosu and Landman (1998) noted that this difference in maximal/non-maximal IHRCs correlates to the presence (Lakhota) or absence (Quechua/Japanese) of an overt D, but this has been challenged by Gur languages, which have definite articles and maximalizing IHRCs (Hiraiwa et al. 2017). A more robust correlation has been claimed to be the correlation between island sensitivity (sensitivity to CNPC) and the maximal interpretation of the internally headed relative clause (Grosu 2002, Watanabe 2004, Hiraiwa et al. 2017).

Lakhota, the language whose internally headed realtive clauses are not island sensitive, also has non-maximalizing internally headed relative clauses. In 13, the interpretation of this internally headed relative clause is not maximal. The set described by the matrix clause, the set of apples that the speaker wants, does not have to contain all the elements of the set described by the relative clause, the set of well-washed apples. This sentence is felicitous even if the speaker only wants one of the well-washed apples.[2]

(13) [[**Thaspą** wąži tąyą yužaža pi] cha] wachį
apple a-IRR well wash PL SM I-want
'I want an apple (nonspecific) that is well washed.'
(**Lakhota**; Williamson 1987, cited in Grosu & Landman 1998: (92a))

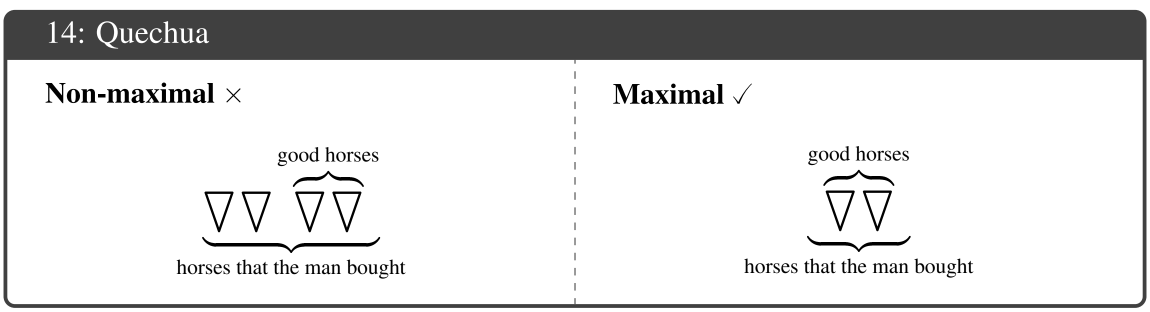
[2]It is not clear from the paper whether a maximal interpretation is allowed in this sentence, so the Maximal interpretation in the figure for 13 is marked as '?'.



13: Lakhota

Quechua is a language whose internally headed relative clauses are island sensitive. The interpretation of its relative clauses has been claimed to be obligatorily maximal, as shown in 14. The set described by the matrix clause, which is the set of horses that are good horses, must have all the same elements as are in the set described by the relative clause, which is the set of two horses that the man bought. It is not possible to use this sentence to say that only some of the horses that the man bought were good horses.

(14) [Nuna **ishkay bestya**-ta ranti-shqa-n] alli bestya-m ka-rqo-n
man two horse-ACC buy-PERF-3 good horse-VAL be-PAST-3
'The two horses that the man bought were good horses.'
Unavailable interpretation: 'Two horses that the man bought were good horses.'
(**Quechua**; Dayal 1991, cited in Grosu & Landman 1998: (93a)



14: Quechua

Japanese is another language with island sensitive internally headed relative clauses. It has also been said to have obligatorily maximal internally headed relative clauses.[3] 15 illustrates this. Here the set described by the matrix clause is the set of cookies brought to the party by Taro. This set must be coextensive with the set of cookies Yoko put in the fridge.

[3]However, there are some, such as Kubota and Smith (2007) who disagree.

(15)

Taro-wa	[[Yoko-ga	reezooko-ni	**kukkii-o**	**hotondo**	irete-oita]-no]-o
Taro-TOP	Yoko-NOM	refrigerator-LOC	cookie-ACC	most	put-AUX-NM-ACC

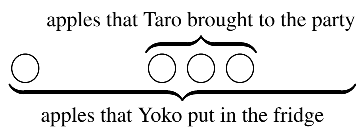
paatii-ni	motte	itta.
party-to	brought	

‘Yoko put most cookies in the refrigerator and Taro brought them, *some to the party.’

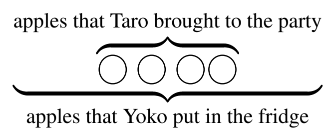
(**Japanese**; Shimoyama 1999, cited in Grosu 2002: (28b)

15: Japanese

Non-maximal ×



Maximal ✓



Thus, Japanese and Quechua seem to have maximalizing, island sensitive IHRCs, and Lakhota seems to have non-maximalizing, island insensitive IHRCs.

Summary

	Maximalizing	Island sensitive
Lakhota	×	×
Quechua, Japanese	✓	✓

4 The case of Yǔn Shan

4.1 Sensitivity to islands

Yǔn Shan internally headed relative clauses do appear to be island sensitive. The Shan equivalent of the Lakhota example in 9 is not acceptable, as shown in 16.

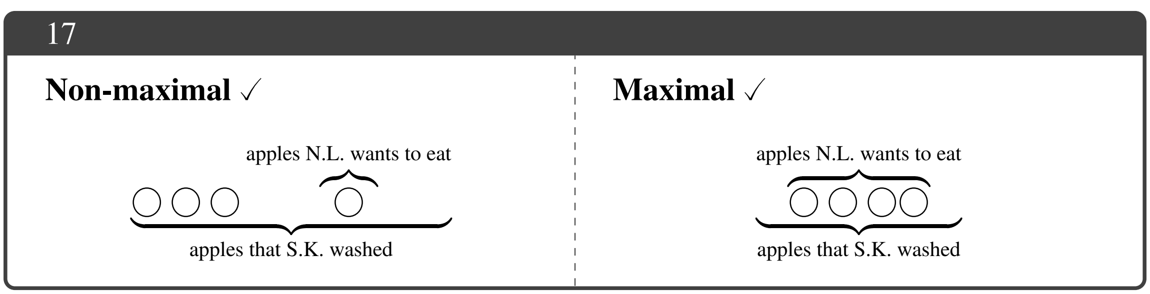
(16) *[ʔăn Nan Li waa kăn táŋhen [ʔăn kón ʔàan **lik** nâj]] mán lĕŋ.
COMP Nan Li spoke together with COMP person read book this 3 red
Intended: 'The book that Nan Li spoke with the people who read (it) is red.'

4.2 Maximality

Given that Shan lacks overt determiners and is sensitive to island constraints, we might expect its IHRCs to be maximalizing. However, its IHRCs appear to allow non-maximal interpretations, giving a counterexample to the correlation between island sensitivity and maximality noted by (Grosu 2002, Watanabe 2004, Hiraiwa et al. 2017).

In 17, the set described by the matrix clause is the set of apples that Nan Li wants to eat, and the set described by the relative clause is the set of apples that Saj Kham will wash. If this were a maximalizing internally headed relative clause, we would expect the follow up sentence, which restricts the number of apples that Nan Li wants to eat to one, to be infelicitous. A maximal interpretation would require Nan Li to want to eat all the apples that Saj Kham will wash.

(17) Nan Li khaj cǐn [ʔăn Saj Kham te lâaŋ **màmô** nâj]. Mán khaj cǐn hwí.
Nan Li want eat COMP Saj Kham will wash apple this 3.SG want eat CL.RND
'Nan Li wants to eat apples that Saj Kham will wash. She wants to eat one.'



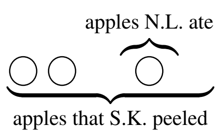
Similarly, 18 shows that a maximal interpretation is not obligatory. When there is a number modifying the internal head of the relative clause, it restricts the number of elements described by the relative clause but not the number of elements described by the matrix clause. The set described by the matrix clause is the set of apples that Nan Li ate, and the set described by the relative clause is the set of apples that Saj Kham peeled. The interpretation of this sentence is that Saj Kham peeled three apples and Nan Li ate some of the peeled apples. However, Nan Li does not have to have eaten all three peeled apples.

(18) Nan Li cǐn pěn [ʔăn Saj Kham pɤ̌k **màmɔ̂ săam hwí** nâj].
Nan Li eat up COMP Saj Kham peel apple 3 CL.RND this
'Nan Li ate apples that Saj Kham peeled which are three in number.'

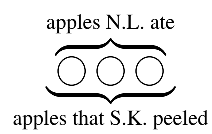
- •Number of apples Saj Kham peeled: 3
- •Apples Nan Li ate: some number of the peeled apples

18

Non-maximal ✓



Maximal ✓

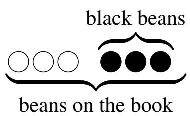


Finally, it is possible to describe elements in a set of things using an internally headed relative clause and then describe elements of the same set of things with contradictory information without it being infelicitous. This is similar to the Consistency test used by Dayal (2004) to identify the true definite determiner of a language. In 19, the internally headed relative clause describes the set of beans on the white book. The matrix clause is the set of black beans. If this were interpreted maximally, we would expect all the beans on the book to be black. However, it is possible to follow up that sentence with one that says that some of the beans are not black. This suggests that a non-maximal interpretation is possible. Importantly, in this context, there are no beans anywhere except on the white book, so it is not possible for 'some beans' to be describing other beans in the context. Also, this sentence is not being interpreted generically in the context. The interpretation is that there are black beans and non-black beans on the white book.

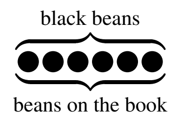
(19) [ʔǎn **thò** mí nɤ̌ lik khǎaw nâj] mán lǎm. **Thò kamphɔŋ** nâj mán mà lǎm.
COMP bean exist on book white this 3.SG black bean some this 3.SG NEG black
'Beans on the white book are black. Some beans are not black.'
(Context: There are beans on the book and nowhere else.)

19

Non-maximal ✓



Maximal ✓



Yǔn Shan does not fit with the typology that connects island sensitivity of IHRCs with a maximal interpretation. It has internally headed relative clauses that are sensitive to island constraints, but it also allows non-maximal interpretations of its internally headed relative clauses.

Summary

	Maximalizing	Island sensitive
Lakhota	×	×
Quechua, Japanese	✓	✓
Shan	×	✓

5 The case of Navajo

Navajo seems like a good candidate to fit with the Shan IHRC type, with one caveat. Grosu (2012) amends his typology by introducing data found in Navajo. Navajo, like Shan, has IHRCs that are non-maximal but sensitive to island constrains, as in 20-21. The English translation of 20 uses an indefinite to modify the head, which is not compatible with a maximal interpretation. 21 has the head *łééchąąí* 'dog' being extracted out of another relative clause with the head *hastiin* 'man'. This is not possible in Navajo, so it is clearly sensitive to the CNPC.

(20) [**Bilasana** hazho'o tanasgis-ígíí] nisin.
apple carefully washed-REL 1-want
'I want an apple that is well washed.' (**Navajo**; Grosu 2012: (4))

(21) *[Hastiin **łééchąąí** bishxash-ę́ę] be'eldooh néidiitá-(n)ę́ę nahał'in.
man dog bit-REL gun pick-up-REL bark
'The dog that the man who was bitten by (it) picked up the gun is barking.'
(**Navajo**; Grosu 2012: (49b))

Strong quantifiers in Navajo seem to obligatorily be interpreted with matrix clause scope, as in 22. If *ałtso* 'all' were interpreted inside of the clause, it would have the interpretation that John bought all the cars (and motorcycles) from Bill.

(22) [John Bill **chidí t'áá ałtso** (**dóó dzi'izí dilchxoshí t'áá ałtso**) yaa
John Bill car 3 all and motor cycle 3 all from
nayiisnii'ę́ę] t'éiyá nizhónígo nidaajeeh.
3.3.buyP.REL only well da.3.run.1
'All the cars (and all the motorcycles) that John bought from Bill —and only those —run well.' (**Navajo**; Grosu 2012: (48))

Grosu analyzes this internally headed relative clause as being a case of 'cyclic re-merger' until the head is outside of the relative clause and claims that the source of island sensitivity is 'traceable to whatever factors require matrix scope for IHs [internal heads]' (Grosu 2012: 25). This suggests that there is a connection between the matrix scope-taking quality of some Navajo quantifiers with the formation of internally headed relative clauses. Grosu notes that this idea fits with Hastings's (2004) claim that Cuzco Quechua IHRCs are restrictive when strongly quantifying and maximalizing otherwise. Hastings (2004) had claimed that strongly quantified heads in Cuzco Quechua internally headed relative clauses obligatorily take matrix scope.

While it would be nice to be able to connect these two observations, it cannot be the whole story: in Yǔn Shan, quantifiers seem to take surface scope obligatorily. This can be seen in the contrast between 23 and 24. In 23 is an externally headed relative clause with a strong quantifier *mɤmót* 'all' modifying the head in the matrix clause. The interpretation of this sentence is that Saj Kham peeled an unspecified number of apples, and Nan Li ate all of those peeled apples. The quantifier takes scope in the matrix clause where it appears. In 24, on the other hand, the quantifier appears inside the relative clause, and the head is either inside or outside the clause. The interpretation when the quantifier is inside the clause is that Saj Kham peeled all the apples (12 in this context), and Nan Li ate some unspecified number of peeled apples. The quantifier takes scope inside the relative clause where it appears. There is a clear difference in meaning that corresponds to the position of the quantifier, not the position of the head.

(23) Nan Li cǐn [**màmɔ̂ mɤmót** [ʔǎn Saj Kham pɤ̀k] ʔǎn nâj].
Nan Li eat apple all COMP Saj Kham peel CL this
'Nan Li ate all the apples that Saj Kham peeled.' (**Shan; Post-Head**)

- Number of apples Saj Kham peeled: some number of apples
- Apples Nan Li ate: all the apples peeled by Saj Kham

(24) Nan Li cǐn pěn [(**màmɔ̂**) [ʔǎn Saj Kham pɤ̀k (**màmɔ̂**) **mɤmót**] nâj].
Nan Li eat up apple COMP Saj Kham peel apple all this
'Nan Li ate apples that Saj Kham peeled all of.' (**Shan; Post-Head/IHRC**)
(Context: there are 12 apples)

- Number of apples Saj Kham peeled: 12
- Apples Nan Li ate: some number of the peeled apples

This data is motivation for an analysis of internally headed relative clauses found in Shan and Navajo where the interpretation is non-maximal but the syntax is island sensitive. While Grosu (2012) has suggested that the analysis could be connected to the matrix scope taking qualities of some quantifiers in Navajo and Cuzco Quechua, that cannot be the whole story since Shan quantifiers obligatorily take scope where they appear. Further, this type of internally headed relative clause is only available in Cuzco Quechua when the head is strongly quantified, but Navajo internally headed relative clauses are all non-maximal/restrictive and island sensitive. It is not clear how an analysis that relies on strongly quantified heads can account for the examples that lack strong quantification.

This data from Yŭn Shan does not fit previous analyses for IHRCs that (1) require a maximal interpretation (Shimoyama 1999, Grosu & Landman 2012, Erlewine & Gould 2016); (2) employ un-selective binding to explain IHRCs that are not subject to wh-island constraints (Watanabe 2004); or (3) rely on strongly quantified heads to explain the availability of non-maximal IHRCs (Grosu 2012).

6 Sketch of Analysis: Head raising

The island sensitivity of IHRCs in Shan, Japanese, and Navajo suggests that there is movement involved, so I will propose LF movement of the head to SpecCP. I will assume a head raising analysis, following the theory proposed Kayne (1994) and Bianchi (1999), and revised by De Vries and colleagues (2002). The head being in SpecCP can explain the island sensitivity of Shan IHRCs. Since it is not necessary to account for maximality, no special analysis is required to account for the interpretation here.

(25) [$_{\text{CP}}$ màmɔ̂$_i$ [$_{\text{C}'}$ ʔăn [$_{\text{IP}}$ [$_{\text{DP}}$ Saj Kham] [$_{\text{VP}}$ [$_{\text{V}'}$ pɤ̌k [$_{\text{ClfP}}$ t$_i$ [$_{\text{ClfP}}$ săam [$_{\text{Clf}'}$ hwí [$_{\text{NP}}$ t$_i$]]]]]]]]]

An interesting thing to consider is how to integrate the analysis for Yŭn Shan IHRCs with the analyses of other island sensitive IHRCs, as are found in Japanese and Navajo, other than simply having a null definite operator appearing where necessary in Japanese and Navajo. The analysis in Grosu & Landman 2012 could be used to account for this data. Grosu and Landman (2012) use event semantics and operator movement instead of covert head movement. The definteness of Japanaese IHRCs is accounted for using a null definiteness operator that is above *-no* in the syntax (Grosu & Landman 2012: 179). Shan could simply lack this null definite operator. Still, finding a more principled explanation for the appearance of a null definte operator is important for future research.

REFERENCES

ALDRIDGE, EDITH. 2004. Internally headed relative clauses in Austronesian languages. *Language and Linguistics* 5.99–129.

BIANCHI, VALENTINA. 1999. *Consequences of antisymmetry: Headed relative clauses*, vol. 46. Walter de Gruyter.

BONNEAU, JOSÉ. 1990. Logical form and an analysis of the matching effect in free relatives. *McGill Working Papers in Linguistics* 6.137–166.

COLE, PETER. 1987. The structure of internally headed relative clauses. *Natural Language & Linguistic Theory* 5.277–302.

DAYAL, VENEETA. 1991. The syntax and semantics of correlatives. *Natural Language & Linguistic Theory* 9.637–686.

DAYAL, VENEETA. 2004. Number marking and (in) definiteness in kind terms. *Linguistics and philosophy* 27.393–450.

DE VRIES, MARK, ET AL. 2002. *The syntax of relativization*, vol. 53. LOT Utrecht.

DRYER, M., and M. HASPELMATH. 2013. *WALS online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: http://wals.info/.

ERLEWINE, MICHAEL YOSHITAKA, and ISAAC GOULD. 2016. Unifying Japanese relative clauses: copy-chains and context-sensitivity. *Glossa* 1.1.

GLICK, I., and S. T. MOENG. 1991. *Shan for English speakers*. Dunwoody Press.

GROSU, ALEXANDER. 2000. The semantic diversity of internally-headed relative clauses. *Naturally*, 143–152.

GROSU, ALEXANDER. 2002. Strange relatives at the interface of two millennia. *Glot International* 6.145–167.

GROSU, ALEXANDER. 2012. Towards a more articulated typology of internally headed relative constructions: The semantics connection. *Language and Linguistics Compass* 6.447–476.

GROSU, ALEXANDER, and FRED LANDMAN. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6.125–170.

GROSU, ALEXANDER, and FRED LANDMAN. 2012. A quantificational disclosure approach to Japanese and Korean internally headed relatives. *Journal of East Asian Linguistics* 21.159–196.

HASTINGS, RACHEL ELIZABETH. 2004. *The syntax and semantics of relativization and quantification: The case of Quechua*. Cornell University dissertation.

HIRAIWA, K.; G. AKANLIG-PARE; S. ATINTONO; A. BODOMO; K. ESSIZEWA; and F. HUDU. 2017. A comparative syntax of internally-headed relative clauses in Gur. *Glossa: A journal of general linguistics* 2.

HIRAIWA, KEN. 2003. Relativization in Buli. *Studies in Buli Grammar. Working Papers on Endangered and Less Familiar Languages* 4.45–84.

KAYNE, RICHARD S. 1994. *The antisymmetry of syntax*. 25. MIT Press.

KUBOTA, YUSUKE, and E ALLYN SMITH. 2007. The japanese internally headed relative clause is not an e-type pronoun. *Proceedings of formal approaches to Japanese linguistics*, vol. 4, 149–60.

ROSS, JOHN ROBERT. 1967. *Constraints on variables in syntax*. Ph.D. dissertation, Massachusetts Institute of Technology.

SHIMOYAMA, JUNKO. 1999. Internally headed relative clauses in Japanese and E-type anaphora. *Journal of East Asian Linguistics* 8.147–182.

WATANABE, AKIRA. 2004. Parametrization of quantificational determiners and head-internal relatives. *Language and Linguistics* 5.59–97.

WILLIAMSON, JANIS. 1987. An indefiniteness restriction for relative clauses in Lakhota. *The Representation of (In) definiteness*, 168–190.

An anaphora-based review of the grammar/pragmatics division of labor

KOICHI NISHIDA

Yamaguchi Prefectural University[*]

1 Introduction

This study attempts to develop Ariel's 2010, 2017 argument that the GRAMMAR/PRAGMATICS DIVISION OF LABOR has to be defined in terms of the code/inference distinction. Ariel (2010: 249) says that 'not everything that is external to the grammar but pertains to language use is pragmatic'. Given this, it leads to the question of how to identify the areas of language use that are external to grammar, but are pragmatic in nature. We need to have a principle on which to identify the labor assignment of pragmatics, as divided from that of grammar and also from those of other fields like sociolinguistics. This article offers one such way with anaphoric definite noun phrases called 'SPECIFYING ANAPHORS' (Skirl 2007), showing that defects in grammar feed pragmatic inferences because the division of labor entails the overlap of the two fields.

This article is organized as follows. Section 2 reviews Ariel's 2010, 2017 proposal for the shift from the extentional definition of pragmatics to its intensional definition in terms of the code/inference distinction. Section 3 deals with the basics of specifying anaphors, with focus on why they are mostly used in subject position. Section 4 discusses the difference between pronouns and specifying anaphors, showing that pronouns belong to grammar, but specifying anaphors belong to pragmatics, and that usage conditions of specifying anaphors help clarify how to divide pragmatics from grammar. Section 5 suggests that proverbs are candidates for the historical change from yesterday's inference to tomorrow's code, pointing out what they have in common with pronouns and specifying anaphors. Section 6 offers concluding remarks.

---

[*] Portions of this article overlap with those of my previous studies (Nishida 2017a, 2017b), but the new data was added and the arguments were completely reworked for the purposes of this article. I am indebted to the following people for their advice, comments, questions and constructive criticism on my presentation and on the manuscript of this article: Minoru Nakau, Yukio Hirose, Yoko Hasegawa, Shigeki Seki, Kristen Sullivan, Satoshi Ota, Hiroaki Konno, Kairi Igarashi, Koji Yoshida, Takanori Demizu, Ayumi Suga, Hiroaki Tanaka, Kyoko Kosaka, Koji Kawahara, Kiyomi Tokuyama, Kiyono Fujinaga, Kiyoko Toratani, Kazuo Yamasaki, Yoshiko Matsumura, Tomoyuki Tsuchiya, Amy Wilson, Natsumi Tanaka and Yuji Ankei. This study was financially supported by the 2017 Yamaguchi Prefectural University grant for research and creation activities, and by MEXT/JSPS KAKENHI Grant Numbers 18K00542 and 18H00680. I am solely responsible for the remaining inadequacies.

2 Theoretical background and empirical issues

It is useful to begin the discussion by reviewing Ariel's grammar/pragmatics division of labor. Instead of defining pragmatics extensionally, i.e., as a list of research objectives like reference, speech acts and implicatures, she proposes to define it intensionally, as a study of the inference derived from the meaning of words rather than of the meaning of words themselves. Ariel (2010: 111) offers the following reason to distinguish between grammar and pragmatics:

> The strongest motivation for distinguishing between codes and inferences is the well-known fact that many of our inferences are innovative uses which could not have already been stored by the addressee. We must distinguish between conventional and innovative uses, between the given and the constructed, … The former are grammatically encoded, the latter pragmatically inferred on the spot.

She argues that grammar and pragmatics are distinguished only by the code/inference distinction, as it captures the difference in the kind of meaning communicated; grammar deals with the meaning encoded and stored in the previously-determined words and phrases, but pragmatics deals with the meaning inferred and newly innovated with the words used in context.

Ariel 2017 argues for a 'case-by-case approach' to the code/inference distinction, because, depending on the case of language use, 'grammar shrinks' and pragmatics 'gets extended', or 'more grammar' is required to use 'less pragmatics'. I argue that it is possible to mediate the border dispute between grammar and pragmatics, with pronouns and specifying anaphors.

3 Basics of specifying anaphors

Skirl (2007: 105) defines specifying anaphors as anaphoric noun phrases 'which contribute specific information about the referent by their semantic content'. They are rich in descriptive content and are used like pronouns to refer back to discourse antecedents, as in 1, where the underlined noun phrase as well as the pronoun is used to refer back to Michelle Obama. In what follows, coreference relations are marked by the underlines for the relevant items.

(1) <u>Michelle Obama</u> is going to meet with Japanese lawyers. {<u>She</u>/<u>The famously fashionable tall woman</u>} will wear blue at the ceremony.

Unlike pronouns, specifying anaphors innovative uses. They are not fixed in form and content, so each of their anaphoric or coreferential readings is not stored in the grammar of previously-determined noun phrases, but instead arises from the inference of the hearer who assumes that the speaker assumes that she, too, adopts Levinson's 2000 coherence-oriented heuristic in 2.

(2) Avoid interpretations that multiply entities referred to.

We discuss specifying anaphors in terms of genre, discourse functions and sentential positions.

First, specifying anaphors are characteristically used in journalistic discourse, as in 3.

(3) There are no flowers or memorials to mark the spot where Tsewang Norbu died. On Aug. 15, the 29-year-old Tibetan monk living in the remote Chinese outpost of Tawu gulped down kerosene, bathed his body in the combustible liquid and struck a match.
(Beech, H. (2011). Burning desire for freedom. *TIME*, 178(19), p.22.)

The descriptive content of a specifying anaphor is specifically made for the referent of its antecedent. This characteristic is extended to apply to the genre of discourse in which specifying anaphors are used. They are used in the context whose content is specifically made for their referents. They are typically used in sports articles, celebrity gossip articles and biographies, all of which focus on specific individuals with unique features, and they are used to refer back to such individuals. The writers in these genres know much more about the topic individuals than most if not all of their readers do. However, specifying anaphors are not used in academic papers or in administrative documents, where pronouns are predominant anaphoric expressions; pronouns have highly generalized descriptive content, and are suited to the context whose content is generalized over other cases besides the one referred to in that context in question. The writers in these genres are supposed to distance themselves from the topic referents, so they need not, and should not, inform the reader of the specific attributes of topic individuals.

Second, specifying anaphors bring new information into discourse, at least for the current hearer. In 3, the readers can understand that the underlined noun phrase is true only of Tsewang Norbu, even if it is entirely new to them. In this respect, Heim (1982) argues that the descriptive content of definites is presupposed, without clarifying by whom it is presupposed. Specifically, Heim (1982: 162) notes that in a sequence like *A dog barked at a cat. The little coward ran away*, *the little coward* refers back to *a cat* without requiring 'a context in which it is presupposed that every cat that was barked at by a dog is a little coward', or without making us assume in this way.

Similarly, specifying anaphors suggest that their descriptive content need not be presupposed, at least by the current hearer. This article tries to answer why and how they can be used in this way.

Umbach (2001: 264-5) argues with the examples in 4 that, like anaphoric pronouns like the one in 4a, most cases of anaphoric definites like *the man* in 4b denote subsuming properties.

(4) a. A man came into the bar. He was carrying a black suitcase.
b. A priest came into the bar. The man was carrying a black suitcase.

Like pronouns, definite noun phrases whose descriptive content is subsuming or semantically general make good anaphoric expressions, because what the speaker presupposes about the referents of such noun phrases can be readily shared by the hearer as well. [1]

Umbach (2001: 265) also offers an apparently inconsistent example, as in 5, which she assumes to be part of a newspaper article, arguing that the description of the underlined noun phrase 'conveys information which cannot be inferred from the previous discourse.'

(5) (newspaper article reporting on a trial: …This morning the court heard the defendant.) … The 34-year-old father of two teenage daughters claimed to be innocent.

Umbach says that the underlined noun phrase in 5 is equivalent to a noun phrase appositive like *the defendant, 34-year-old father of two teenage daughters*, and that the information added in anaphoric noun phrases like this 'is presented as if it were already known'.

Umbach's argument reminds us of Prince's 1978 classification of *it*-clefts into the stressed focus type and the informative-presupposition type. The former has a *wh*- or *that*-clause (cleft clause) whose content is presupposed by the hearer, as Prince (1978: 896) exemplifies with 6.

(6) So I learned to sew books. They are really good books. It was just the covers that are rotten.

In 6, the fact that something is rotten is known from the first sentence, so the cleft clause of the third sentence carries the presupposition the speaker shares with the hearer. Cleft clauses of this type are close to anaphoric definites whose content is known to the hearer, like the one in 4b.

[1] These observations lead Levinson (2000: 270) to say that 'the more semantically general an expression is, the greater the likelihood that it will be descriptively inadequate for reference identification', and so is likely to be anaphoric. He regards null subjects as ideal anaphoric expressions, because they are maximally general, or semantically null. It is clear, however, that specifying anaphors have rich descriptive content, and cannot be explained away in terms of semantic generality.

The informative-presupposition type has a cleft clause carrying new information, as Prince 1978: 898, 902 illustrates with 7a and 7b, where ## stands for a discourse initial position.

(7) a. ##It was just about 50 years ago that Henry Ford gave us the weekend. On September 25, 1926, in a somewhat shocking move for that time, he decided to establish a 40-hour work week, giving his employees two days off instead of one.

b. But why is the topic so important? Apparently, it is the topic that enables the listener to compute the intended antecedents of each sentence in the paragraph.

To cite Prince 1978: 899-900, *it*-clefts like those in 7 mark the cleft clause 'as fact, known to some people, although not yet known to the intended hearer' and occur in 'persuasive discourse'.

Lambrecht (1994: 70-1) argues that the two types of *it*-clefts reflect two kinds of speaker-hearer relations. The stressed focus type is the original type, and the content of its cleft clause belongs to the common ground between speaker and hearer. The informative-presupposition type is an extension of the original, and its cleft clause carries a piece of information which is shared by the community of speakers and their peers, to which the current hearer is an outsider.

Lambrecht's comments apply to the two types of anaphoric definites, those whose content is presupposed and specifying anaphors. The former are used in homogeneous speaker-hearer communities whose members share the presupposition about antecedent referents. The latter are used in heterogeneous communities in which some members do not share the presupposition that others have, and they have to accept it as being presupposed about the antecedent even if it is new for them. Thus, speakers can use specifying anaphors as part of their AUDIENCE DESIGN (Clark 1996), by which they treat the current hearers as lacking knowledge about the topic, as is often the case with journalists who persuade readers of the correctness of their stories.[2] Thus, specifying anaphors work with discourse organization rather than with sentence grammar.

[2] Declerck (1988: 224) classifies Prince's informative-presupposition *it*-clefts into two subtypes, discontinuous clefts like the one in 7a and unaccented-anaphoric focus clefts like the one in 7b. In the former, both the focus and the cleft clause are new; in the latter, the focus is old, and the cleft clause is new but represented as old. Specifying anaphors are closer to the latter than to the former. Functionally, *it*-clefts and definites are similar. Just as *it*-clefts are classified according to whether or not their cleft clause carries information shared by both speaker and hearer, definites are classified according to whether or not their descriptive content is familiar to both speaker and hearer. In light of the code/inference distinction, however, *it*-clefts and definites are different. As shown by Prince 1978 and Declerck 1988, the classification of *it*-clefts involves a number of syntactic correlates: for example, the stressed focus type, but not the informative-presupposition type, allow *that* or a *wh*-word to be deleted

Next, I collected over 600 examples of specifying anaphors, and classified them according to their positions in the sentence, as in table 8. I found that of 505 examples in subject, 356 are main clause subjects and have their antecedents in main-clause subject, too (Nishida 2017a, b).

(8) The sentential positions of specifying anaphors

Subjects	Possessives (Those in subject)	With prepositions	Objects	Sum
505	44 (29)	29	22	600

For the present purposes, we focus on the fact that specifying anaphors are used by far more frequently in subject. This frequency is of pragmatic nature, since they are not grammatically excluded in other positions. Their subject-orientation is functionally reasonable; it is an outcome of the fact that they have more to do with discourse organization than with sentence structure.

Thanks to their dependency on other referring expressions in the same context, anaphoric expressions can perform two contrastive functions. The first is a grammatical function of giving a semantic role like possessor or patient to the referent of their antecedent in the sentence to which it belongs. This function is particularly suited to case-marked pronouns like *his* and *him*, since they express semantic roles that its referent has in a word. The second is an informative function of providing hearers with additional information about their antecedent referents. This is performed by noun phrases whose descriptive content can bring new information to hearers.

The two functions cannot be performed by a single anaphoric expression, as stated in 9.

(9) The function of an anaphoric expression may be either grammatical or informative.

As stated with *may*, this generalization correctly allows that anaphoric expressions may be freed from both functions, which is exemplified by null subjects in elliptical sentences.[3]

As a piece of evidence for the generalization in 9, in 10, possessive pronouns like *his*, but not possessive specifying anaphors like *the smart 29-year-old sociologist's*, can accompany head nouns that refer to separate entities from their possessors, like *(one's) colleagues*.

---

from the cleft clause. However, the classification of definites has no such correlates. This suggests that the classification of *it*-clefts belongs to grammar, but that of definites to pragmatics.

[3] In making 9, I benefited from Lambrecht (1994: 185), who proposes the Principle of the Separation of Reference and Role, which is summarized as 'Do not introduce a referent and talk about it in the same clause'. On this principle, reference to a discourse referent is carried out by noun phrases, and its role assignment by pronouns. I extend his principle, saying that pronouns and specifying anaphors have contrastive functions, so they exhibit different distributions both in sentence and in discourse.

(10) Dr. Richard Brown published a monograph, because

a. his colleagues introduced the 29 year old smart sociologist to the publisher.

b.??the 29 year old smart sociologist's colleagues introduced him to the publisher.

Anaphoric possessives are grammatically important, because they link the referents of their antecedents to another referent of a nominal they place in the head of an entire noun phrase. By 9, grammatically important anaphoric expressions are primarily filled by pronouns which are not important for the informative function. The unacceptability of 10b illustrates that there are no anaphoric expressions that are important in both grammatical and informative functions.

As far as I have observed, most possessive specifying anaphors occur with subjects, as in table 8, and all of them accompany head nouns that represent parts of the referents of their own antecedents rather than other separate entities, although such examples are not frequent. For example, *(one's) taste* in 11 is a part of the possessor referred back to by *the media baron*.

(11) Silvio Berlusconi ... is not known for his restraint. The media baron's taste for the high life, his dalliances with young women, and his lazy moral compass are now legendary.
(NewsBeast. (2011, August 22&29). *Newsweek*, p.16.)

Possessive specifying anaphors can be used if they precede nouns denoting parts of the possessor that are not counted as separate entities, and thus are exempted from the grammatically important function of linking its antecedent referent to a separate entity in a noun phrase.

Anaphoric expressions that are important in the informative function are primarily filled by descriptive noun phrases in subject, because subjects are grammatically the least important in sentence structure; specifying anaphors mostly occur where they have the least to do with the structure of the sentences to which they belong. Moreover, the subject is an unmarked position of the topic which is a continuation of the previous discourse (Lambrecht 1994, Chapter 4), so specifying anaphors in subject are suited to providing hearers with new or additional information about the topic whose identity of reference belongs to the old information obtainable from the previous discourse. The new information added to the old can be as rich as it is in 12.

(12) Ralf Möller greets her at the Reckling-hausen podium. The 6-ft. 5-in. (196cm) former champion bodybuilder and *Gladiator* actor (he played Hagen, a Germanic barbarian, in the

Oscar-winning epic), who was born in the city and these days is literally big in Hollywood, has flown from Los Angeles to lend his muscle to Merkel's cause.
(Mayer, C. (2013). Angela Merkel's unfinished business. *TIME*, 182(13), p.22.)

Common types of additional information expressed by specifying anaphors include attributes of individuals like age, nationality, physical properties, social and family status, occupation and specialty, and when combined, these attributes apply only to specific individuals, as in 3 and 12.

4 The difference between pronouns and specifying anaphors

This section shows that pronouns and specifying anaphors have different distributions, but specifying anaphors rely on the grammar of pronouns to determine where they may occur. While pronouns are used both in sentence anaphora and in discourse anaphora, specifying anaphors are, in principle, excluded from sentence anaphora, as shown by the contrast in sentence 13.

(13) Michelle Obama said that {she/??the famously fashionable tall woman} would meet with Japanese lawyers.

In 13, only *she* is fine. Pronouns like this are given the LOGOPHORIC reading via a grammatical convention with two requirements; (i) they have to be in complements of verbs of saying and thinking, and (ii) they have to be coreferential with matrix subjects (Huang 2000: 181). These two conspire to represent the matrix subject in 13 as the original speaker of a first-person utterance like 'I will meet with Japanese lawyers' (Kuno 1972, Vandelanotte 2004).

The logophoric reading of pronouns belongs to grammar, because it is stored in the previously-determined words of pronouns, and does not apply to specifying anaphors. They are not pronouns, and so are grammatically disallowed to stand for the original speaker in 13.

However, when original speakers in matrix subject are replaced with their spokespersons, as in 14, the second requirement is loosened and the convention is made defective.

(14) Michelle Obama's secretary said that {she/the famously fashionable tall woman} would meet with Japanese lawyers.

Here the complement can no longer give a logophoric reading to the pronoun grammatically, but instead, it accepts a specifying anaphor. As the coherence-oriented heuristic in 2 remains valid in the defective convention to give a coreferential reading to the definite noun phrase in 14 in the same way as it does in contexts like 1 and 3, showing that grammar and pragmatics overlap in work with grammatical conventions in different but interdependent ways. [4]

It is thus necessary to understand the grammar/pragmatics division of labor as dividing the labor assignment of pragmatics from that of grammar in the way stated in 15.

(15) Pragmatics works with grammatical conventions where grammar is defective.

It follows from 15 that the division of labor means to provide defective grammatical codes with inferences to make grammatical conventions usable in discourse.[5]

To see how the principle in 15 works, consider the following example where Feng Tianwei is a Singaporean table tennis player and Jing Junhong is her coach; Jing spoke in her place in the interview after the game, and the complement of *maintain* introduces a specifying anaphor.

(16) Feng fell 3-4 … to Hong Kong's Lee Ho Ching to end the tournament in a disappointing sixth place. Despite the underwhelming finish, Jing maintains that the 27-year-old has not under-performed.

(Toh Ting Wei. (2014, March 23). Yu shines, Feng falters. *the sundaytimes*, p.52.)

---

[4] Koji Kawahara commented on my presentation at *BLS* 44, saying that the contrast between 13 and 14 may come from structural differences between the two. More structural arguments are needed to clarify the difference between the antecedent in matrix subject and that in the possessive of the matrix subject. Rather, the current concern is with showing that distribution and interpretation of pronouns are more explicitly stated in structural terms than those of specifying anaphors. He also pointed out that grammatical defectivity is a matter of degree. It may be that conventions for reference assignment of pronouns are partially defective, but remain in grammar; those for specifying anaphors are more defective, and belong to pragmatics; those for proper names are still more defective, and go beyond the realm of pragmatics, since reference to proper names builds on the world knowledge of speaker and hearer rather than their linguistic knowledge. The degree of defectivity is also seen from the fact that, depending on the type of descriptive content, definites range from those that behave like pronouns to those that behave like proper names. Dubinsky and Hamilton (1998:687-8) show that epithets like *the idiot* resemble pronouns, but definites denoting titles like *the president* do not, as illustrated by the contrast in *John ran over a man who was trying to give {the idiot/*the president} directions*. As argued in Section 5, epithets may be a model of analogous reading of proverbs in that they refer back to, and evaluate, their antecedent referents at the same time. Examples of the latter resemblance include definite common noun phrases that are used as proper names or place names in reference to unique entities as like *the Pentagon*, *the United Nations* and *the Great Basin* (Lyons 1999, Chapter 1).

[5] In making 15, I benefited much from Iwata 2012, Chapter 6, who compares grammatical constructions like double object constructions and expression variants like echo-questions. He notes that the former, but not the latter, have specific sentence forms, and argues that the two aspects of language require two different kinds of theories; the former are to be explained in terms of combination of verb semantics and constructional analysis, i.e. in terms of grammar, and the latter in terms of pragmatics. Iwata's division of grammar and pragmatics is similar both in spirit and content to that of Ariel's.

It is impossible to derive *the 27-year-old* in 16 from the term with which Jing addresses Feng nor from her self-reference, because reference to a person's age is normally avoided in the terms with which to address that person, and in that person's self-reference. Rather, it is reasonable to assume that *the 27-year old* comes from the words of the reporter of this article. This assumption is supported by the following pair of the same news appearing in two different newspapers.

(17) a. Jagger's spokesman said the British rocker was 'completely shocked and devastated' at the news while his model daughter Georgia May, 22, pulled out of Melbourne Fashion Week.

(Rolling Stones frontman Mick Jagger stunned by designer girlfriend's death. (2014, March 19). *The Straitstimes*, p. C10.)

b. A spokesman for Mick Jagger said that the singer was completely shocked and devastated by the news.[6]

The specifying anaphors in 17a and 17b mean that the reporter, but not Jagger's spokesman, is responsible for their descriptive content; the two articles were written by two different reporters, and the two definites were employed. To use a term of Mey 1999: 128, the reporter comes close to a narrator in having the power of 'a near-divine omniscience' about the topic, and can control much more information than the original speaker and the spokesperson can (Coulmas 1985).

A question arises as to where the reporter as a narrator can show his power of omniscience about the topic. He has to follow the grammatical convention for logophoric pronouns when the original speaker is in matrix subject and he refers back to that speaker in the complement. By contrast, he is free from it when the original speaker is not in matrix subject, since there is no such grammatical convention for the spokesperson's complement of verbs of saying. This is the niche where he can use definite noun phrases as well as pronouns for anaphoric expressions.

With 9, if the reporter as a narrator uses anaphoric expressions to show omniscience about the topic referent, he has to use them in syntactic positions which are free from the grammatical function; his omniscience is expressed through the informative function. Thus, the specifying anaphors in 16 and 17 belong to the reporter's discourse where grammar is defective.

5 Proverbs and the overlap of grammar and pragmatics

[6] See: http://edition.cnn.com/2014/03/17/showbiz/celebrity-news-gossip/lwren-scott-designer-obit/

Finally, I argue that the interdependency of grammar and pragmatics in 15 is responsible for the historical change from ‘yesterday’s inference’ to ‘tomorrow’s code’ (Ariel 2010: 273). Proverbs in discourse are case in point. In 18, *one swallow* is taken to be an analogy to what the student in question had with the style expressed in the first sentence.

(18) One student complained that she did not like the style of a particular person who was her inaugural ‘baptism’ lecturer. The wise response came from another tutor that one swallow does not make a summer.
(McIlroy, D. (2003). *Studying at university: How to be a successful student*. Sage. p.37.)

Outside of the proverb, *one swallow* will never invite such a reading, because the proverb as a whole is an analogy to the preceding context. Every proverb accompanies analogous readings, and the question is whether they are a matter of grammar, or that of pragmatics. I suggest that they have much in common with discourse anaphora, especially with specifying anaphors.

Proverbs in discourse may have started with inferences on the spot for analogies, and on this stage they resemble specifying anaphors. They follow the heuristic in 2, too. They do not multiply situations to talk about; they just express analogies to the current situation. On the next stage, they assimilate to anaphoric pronouns. Since proverbs are fixed in form and limited in numbers, they come to function as previously-determined expressions in grammar, and encode a discourse function to let the hearer make an analogy to the situations currently described.

Proverbs are used like anaphoric expressions in discourse, because they share the concept of familiarity with pronouns and definites (Lyons 1999, Chapter 7). Because of their definiteness, pronouns express something familiar to or presupposed by speaker and hearer, and definites express what the speaker assumes to be familiar to, or presupposed by, the hearer. Proverbs, too, express familiarity through their propositional content; they must represent familiar situations in which both speaker and hearer can find something analogous to the current situation.

There is a question of whether analogous readings of proverbs belong to inference, or to grammar. The first choice is supported by the fact that some proverbs allow formal variations and yet keep the same analogous reading. For example, *empty vessels make (the) most noise* has several different versions like *empty vessels make {the most/the loudest} sound*, *empty vessels make {the loudest noise/the greatest sound}, empty {barrels/cans/pots} make the most noise*, and so on, all of which are equally used to give a low value to someone who speaks much, but is not intelligent. This shows that like conversational implicatures, the analogous reading is associated with the meaning, rather than the form, of the proverb, and is a matter of inference.

Proverbs reflect cultural conventions in content, but differ from other cultural conventions like greetings and sociolects in motivating hearers to infer what speakers mean by saying them, i.e., to infer what analogies they make on topic referents. As greetings and sociolects do not invite such inferences, they are matters of sociolinguistics, but not matters of pragmatics.

The first choice is also supported by the similarity between proverbs and definites, and the difference between proverbs and pronouns, with respect to their interaction with inference. Consider the following examples taken from Heim 1982: 21, and Roberts 2003: 335.

(19) a. I dropped ten marbles and found only nine of them. {The missing marble/?It} is probably under the sofa.

b. I dropped ten marbles and found all of them, except for <u>one</u>. <u>It</u> is probably under the sofa.

Definites, but not pronouns, are felicitously used when their antecedents remain implicit, as in 19a. Pronouns, by contrast, need explicit antecedents, as in 19b. A similar observation can be made in relation to the proverb in 20. This is a title of a medical article in which *too many cooks* is a negative analogy to two or more physicians involved in the medication of an elderly patient, although they are only indirectly mentioned as a part of the premodifier of *involvement*.

(20) Tamblyn, R. M. et al. 1996. Do too many cooks spoil the broth? Multiple physician involvement in medical management of elderly patients and potentially inappropriate drug combinations. *Canadian Medical Association Journal* 154. 1177-84.

This illustrates that like definites, proverbs do not have to have an explicit mention of discourse referents to which they give analogies, and that they interact with the hearer's inference to be linked with such referents as definites do (Akmajian et al. 2010: 387).

The second choice is supported by the fact that proverbs resemble case-marked pronouns in giving semantic roles to the topic discourse referents and to other related referents. In 21, *sleeping dogs* is taken to be an analogy to the manuscripts found at Nag Hammadi.

(21) But centuries later, the manuscripts found at Nag Hammadi present data that conflicts with the lore the early Catholic Church had already established as the basis of their religion. The

current Church couldn't refute centuries of dogma, so they thought it best to let sleeping dogs lie, and the new stuff never got added.
(Smith, K. (1999). *Dogma: A screenplay*. Grove Press. p.108.)

The analogous reading is grammatically constrained: the proverb in 22 analogously identify its null subject as the matrix subject, *they*, i.e. people in the current Church, and its object as the patient affected by them, i.e. the manuscripts in question. Proverbs have nominal arguments that perform the grammatical function as case-marked pronouns do, and follow the either-or choice in 9. They may be assigned the grammatical function as they do not have the informative function that specifying anaphors have. Despite their propositional content, proverbs do not inform the hearer of anything specific about the topic referent. Their literal meanings do not make much sense, but rather assign either a positive or a negative evaluation to the situations they represent. This function is of pragmatic origin, and thus is expected to be valid across languages.

The second choice is likely to be supported by the cross-linguistic availability of analogous readings of proverbs. They are also available with Japanese proverbs like *inu-mo aruke-ba bou-ni ataru*, which is literally translated as 'when a dog walks around, it hits a bar, too'.

According to Shogakukan 2000, this proverb was originally used in two contrastive senses in the 18th century. The first sense means that people should not walk around without a specific purpose, because they face trouble just like a dog which walks around and hits a bar. The second sense means that people should walk around even when they do not have a specific purpose, because going outside gives them a chance of finding something new. Currently, the second sense is predominant, and goes out of the realm of pragmatic inference which varies by the context, and enters into the realm of grammar where it is specialized to perform a discourse function of giving a positive evaluation to the preceding discourse referent.

(22) Hima tsubushi-ni hon-ya ni i-tta-ra yomi-takat-ta hon-ga
spare-time kill-to bookstore to go-PAST-then read-want-PAST book-NOM
'When I went to a bookstore to kill time, I was able to buy a book that I wanted to read.'
ka-e-ta. Inu-mo aruke-ba bou-ni ataru koto-ga ari-masu.
buy-ABLE-PAST dog-too walk-if bar to hit thing-NOM exist-POLITE
'There are cases when a dog walks around, it hits a bar, too.'

In 22, the implicit subject of the first sentence, i.e. the speaker of 22, is given an analogy to a dog, *inu,* and *bou* 'bar,' makes an analogy to the book that the speaker found by chance. Because of the grammatical function of giving semantic roles to the relevant discourse referents, the subject of the proverb assigns an agent role to the first-person speaker, and the object, or the verb phrase, of the proverb assigns a theme role to the object of the first sentence, thereby evaluating it as a piece of the chance that the speaker was given when he was walking.

Proverbs range over pragmatics and grammar, and are in the shift from inference to code, i.e. a shift in progress where the pragmatics of proverbs and their grammar are equally defective.[7]

6 Conclusion

I have shown that Ariel's division of labor requires the principle in 15 to identify the labor assignment of pragmatics as divided from that of grammar. The grammar/pragmatics division of labor does not mean to assign separate tasks to each field. Rather, pragmatics works with defective grammatical conventions, and specifying anaphors are good examples where it works in this way: they are used where pronouns are not chosen for grammatical reasons. The division of grammar and pragmatics is not always clear, because of the historical change from yesterday's inference to tomorrow's code. This is the case with the analogous readings of proverbs, which are located in history between pragmatic inference for specifying anaphors and grammatical codes for pronouns. Much more studies on defective grammatical conventions, besides those for anaphoric expressions, are necessary to implement Ariel's intensional definition of pragmatics, so as to identify areas of language use where grammar ends and pragmatics starts.

REFERENCES

AKMAJIAN, ADRAIN; RICHARD A. DEMERS; ANN K. FARMER AND ROBERT M. HARNISH. 2010. *Linguistics: An introduction to language and communication*. Sixth ed. Cambridge, Mass.: The MIT Press.

ARIEL, MIRA. 2010. *Defining pragmatics*. Cambridge: Cambridge University Press.

[7] Hedberg (2000) and Gundel et al. (2001) propose that *it*-clefts have definite-noun-phrase equivalents that consist of the discontinuous combination of *it* and cleft clauses in which *it* works as a definite article and the clauses provide descriptive content. In view of the similarities in discourse, proverbs may be equivalent to semantically general definites, too, in that their fixed forms function as a kind of definiteness, and their propositional content is presupposed to be familiar to the relevant speaker-hearer community. As suggested note 4, proverbs resemble epithets, since proverbs and epithets are both semantically general and evaluative. Like proverbs, epithets are located between pronouns and definites, since they are formally definites, but behave like pronouns. I have to leave this issue open here.

ARIEL, MIRA. 2017. Pragmatics and grammar: more pragmatics or more grammar. *The Oxford handbook of pragmatics*, ed. by Yan Huang, 473-92. Oxford: Oxford University Press.

CLARK, HERBERT H. 1993. *Arenas of language use*. Chicago: University of Chicago Press.

COULMAS, FLORIAN. 1985. Direct and indirect speech: general problems and problems of Japanese. *Journal of Pragmatics* 9. 41-63.

DECLERCK, RENAAT. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press, Dordrecht: Foris.

DUBINSKY, STANLEY AND ROBERT HAMILTON. 1998. Epithets as antilogophoric pronouns. *Linguistic Inquiry* 29. 685-93.

GUNDEL, JEANETTE K.; NANCY HEDBERG AND RON ZACHARSKI. 2001. Definite descriptions and cognitive status in English: why accommodation is unnecessary. *English Language & Linguistics* 5. 273-95.

HEDBERG, NANCY. 2000. The referential status of clefts. *Language* 76. 891-920.

HEIM, IRENE. 1982. *The semantics of definite and indefinite noun phrases*. Amherst: University of Massachusetts dissertation.

HUANG, YAN. 2000. *Anaphora: A cross-linguistic approach*. Oxford: Oxford University Press.

IWATA, SEIZI. 2012. (in Japanese) *Eigo-no sikumi-to bumpo-no karakuri* (*The system of English and the mechanism of grammar*). Tokyo: Kaitakusha.

KUNO, SUSUMU. 1972. Pronominalization, reflexivization, and direct discourse. *Linguistic Inquiry* 3. 161-95.

LAMBRECHT, KNUD. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

LEVINSON, STEPHEN C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: The MIT Press.

LYONS, CHRISTOPHER. 1999. *Definiteness*. Cambridge: Cambridge University Press.

MEY, JACOB L. 1999. *When voices clash: A study in literary pragmatics*. Berlin: Mouton de Gruyter.

NISHIDA, KOICHI. 2017a. Two types of pragmatic heuristics and anaphoric expressions. Paper presented at 15th International Pragmatics Conference, Belfast, Northern Ireland.

NISHIDA, KOICHI. 2017b. (in Japanese) Eigo-no kijutu-naiyo-no yutakana meisiku-no doitu-siji-yoho-to taiwa-no teisiki-ka (The coreferential use of English noun phrases that are rich in descriptive content and formulation of dialogue). *Eigo-goho-bunpo kenkyu* (*Journal of English grammar and usage*) 24. 87-102.

PRINCE, ELLEN F. 1978. A comparison of *wh*-clefts and *it*-clefts in discourse. *Language* 54. 883-906.

ROBERTS, CRAIGE. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26. 287-350.

SHOGAKUKAN. 2000. (in Japanese) *Nihon kokugo daijiten* (*The grand dictionary of the Japanese language*). Second ed. Tokyo: Shogakukan.

SKIRL, HELGE. 2007. Metaphorical anaphors: a phenomenon of the semantics-pragmatics interface. *Anaphors in text: Cognitive, formal and applied approaches to anaphoric reference*, ed. by Monika Schwarz-Friesel; Manfred Consten and Mareile Knees, 103-20. Amsterdam: John Benjamins.

UMBACH, CARLA. 2001. (De)accenting definite descriptions. *Theoretical Linguistics* 27. 251-80.

VANDELANOTTE, LIEVEN. 2004. Deixis and grounding in speech and thought representation. *Journal of Pragmatics* 36. 489-520.

Underspecification and scope: The case of gapping

SANG-HEE PARK
*University at Buffalo**

1 Introduction

The syntax and semantics of gapping have been a source of much debate in linguistic theory. Central to this debate is the scope ambiguity first noted by Siegel 1984, exemplified by 1.

(1) Max can't live in San Francisco and Claire in Buffalo.

This sentence has a DISTRIBUTIVE-SCOPE reading in which the negated modal contributed by the auxiliary *can't* is part of each conjunct's interpretation, as in 2a. But the sentence also has a WIDE-SCOPE reading in which the negated modal outscopes the conjunction, as in 2b.

(2) a. $\neg\Diamond$(live-in(Max,SF))$\wedge\neg\Diamond$(live-in(Claire,Buffalo))
... because Max hates humid weather and Clair hates cold winters.

b. $\neg\Diamond$[(live-in(Max,SF))$\wedge$(live-in(Claire,Buffalo))]
... because they can't live apart from each other.

Note that this latter, wide-scope reading is unavailable if gapping does not occur. The ungapped counterpart of 1, given in 3, only has the distributive-scope reading.

(3) Max can't live in San Francisco and Claire can't live in Buffalo.

Wide-scope readings like the one in 1 pose a challenge to the syntax-semantics interface, as there is a mismatch between the surface position of the auxiliary and where it is interpreted. Previous studies propose to account for this mismatch by positing small coordinate structures from which the wide-scope reading can be generated (Coppock 2001;

*I thank Rui P. Chaves, Jean-Pierre Koenig, Hanno Beck, and members of the Syntax-Semantics Research Group at the University at Buffalo for their comments and discussions that contributed to this work. I would also like to thank the audience at BLS 44 for their helpful comments on the presentation that preceded this paper. Any remaining errors are my own.

Johnson 2009; Lin 2002; Toosarvandani 2013). More recently, this approach has led to an analysis in which the ambiguity in 1 is reduced to coordination at two different sites (Potter 2014; Potter *et al.* 2017). However, these analyses all fail to provide an empirically adequate account of gapping.

In this paper, I propose an alternative, semantic analysis in which the various readings of gapping sentences result from different ways of resolving a single, underspecified meaning (Copestake *et al.* 2005; Egg *et al.* 2001; Reyle 1993; Richter & Sailer 2004). The underspecification-based analysis proposed in this work avoids having to postulate a syntactic ambiguity to account for the various readings of gapping, thus offering a simpler account for the interaction between gapping and scope.

2 Previous syntactic analyses

Coppock (2001) and others (Johnson 2009; Lin 2002; Toosarvandani 2013) propose an analysis in which gapping sentences derive from coordination of roughly vP-sized conjuncts. In this analysis, the sentence in 1 derives from the underlying structure in 4.[1]

(4) [$_{\mathrm{T}}$ can't [$_{\mathrm{vP}}$[$_{\mathrm{vP}}$ Max live in San Francisco] and [$_{\mathrm{vP}}$ Claire live in Buffalo]]]

Since the auxiliary is located outside the coordinate structure in this representation, the wide-scope reading is correctly obtained. But assuming this is the only possible syntactic structure for 1, the availability of the distributive-scope reading in 2b is completely unexpected.[2]

One may hypothesize that a vP coordinate structure like the one in 4 is correct for the wide-scope reading but that there is another syntactic structure that can be assigned to 1 that would give rise to the distributive-scope reading. Potter 2014, and more recently,

[1] The surface structure is then derived via a series of movement operations including remnant movement and asymmetric movement of the subject in the first conjunct, which I omit due to considerations of space.

[2] Further contraindications to gapping as vP coordination include cases that involve elements that are moved to or base-generated in the left-periphery of CP, such as topicalized elements (as in (i)) and *why* (as in (ii)).

(i) At our house, we play poker and at Betsy's house bridge. (Sag 1976:265)

(ii) Why did John go by train and why Mary by car? (Repp 2009:34)

Potter *et al.* 2017 take precisely this tack, arguing that the two readings of gapping sentences follow from there being two different syntactic structures for those sentences: vP and CP coordination. In Potter *et al.*'s analysis, the wide- and distributive-scope readings of 1 result from roughly 5a and 5b, respectively (where strikethrough indicates non-pronunciation).[3]

(5) a. Max$_j$ [$_{\mathsf{T}}$ can't [$_{\mathsf{vP}}$ [$_{\mathsf{vP}}$ t_j live in San Francisco] and [$_{\mathsf{vP}}$ Claire$_k$ [in Buffalo]$_l$ ~~t_k live t_l~~]]] $\neg\Diamond(P \wedge Q)$ reading

b. [$_{\mathsf{CP}}$ [$_{\mathsf{CP}}$ Max can't live in San Francisco] and [$_{\mathsf{CP}}$ Claire$_k$ [in Buffalo]$_l$ ~~t_k can't live t_l~~]] $\neg\Diamond P \wedge \neg\Diamond Q$ reading

Crucially, this analysis predicts a tight connection between the site of coordination (vP vs. CP) and interpretation (wide vs. distributive-scope readings). That is, a wide-scope reading results if vP coordination is involved, as there is only a single T above the coordinate structure (as in 5a); a distributive-scope reading results if CP coordination is involved, as each conjunct contains its own T (as in 5b). However, these predictions do not square well with empirical data. A case in point is provided in 6.

(6) A1: Was your father in a bad mood last night?

B: Why? Did he do something strange?

[3] Potter *et al.* pursue a move-and-elide approach, where the remnants evacuate the target of ellipsis via multiple leftward movement. However, it remains to be clarified precisely what derives such movement and why the movement cannot apply to the ungapped counterparts of gapping sentences.

(i) *Max can't live in San Francisco and Claire$_k$ [in Buffalo]$_l$ t_k can't live t_l.

Potter *et al.* (2017:1137–8) allude to an idea along the lines suggested in Pesetsky 1997:342 as a way of licensing leftward remnant movement – namely, the remnants, which introduce novel, unrecoverable information, must be fronted to survive ellipsis. Such an assumption might then be taken to be responsible for preventing the movement in (i).

Crucially, for this argument to hold, multiple leftward movement should not be able to occur without ellipsis accompaniment. But this assumption is undermined by examples such as (ii) (see also Rochemont & Culicover 1990:176 and examples therein).

(ii) A: Which of these books do you think I should give to Leslie?
B: I don't know, but [to MARY] [this book] you should give.

In (iiB) multiple leftward movement has taken place from a constituent that is not targeted by ellipsis. Hence appealing to the notion of recoverability to explain leftward remnant movement is insufficient.

A2: During dinner he **didn't** address his colleagues from Stuttgart or at any time his boss, for that matter. (López & Winkler 2003:241)[4]

On the generally accepted assumption that topicalization in English targets the left edge of CP, a vP-coordination parse is ruled out for 6A2. But this wrongly predicts that a wide-scope reading is unavailable in 6A2, contrary to fact. Native speakers report that 6A2 is understood as 'it is not the case that during dinner the father addressed his colleagues from Stuttgart or that at any time he addressed his boss.' This reading has a structure like $\neg(P \vee Q)$, which is equivalent to the $\neg P \wedge \neg Q$ reading of 7.

(7) He **didn't** address his colleagues from Stuttgart during dinner **and** he **didn't** address his boss at any time, for that matter.

Note that the topicalized NPI *at any time* in 6A2 indicates that this is a case of negation taking wide-scope over both disjuncts (López & Winkler 2003:241), and not a case of distributive-scope of negation under the inclusive interpretation of disjunction. The latter possibility is untenable given the ungrammaticality of 8, which shows that a clause-mate negation cannot license a fronted NPI.

(8) *During dinner he didn't address his colleagues from Stuttgart or **at any time** he didn't address his boss.

Furthermore, there are problems that pertain to the vP coordination component of the analysis of Potter *et al.*. Here, I present the counterevidence from the distribution of *merely* as a case in point (The argument here is adopted from Kubota & Levine 2016:125.[5]) Given that *merely* is a strictly vP modifier, Potter *et al.* would predict that it can appear conjunct-initially in the second conjunct when vP coordination is involved (see 5a). But the prediction fails to hold: Examples like 9 are bad under either a CP or vP coordination parse

[4]Assuming Johnson's (2009) vP coordination analysis to be correct, López & Winkler consider 6A2 as *prima facie* evidence for a vP-internal landing site for topicalization. But why this landing site is available only in the context of gapping is not explained, and I therefore remain skeptical of such an account.

[5]Kubota & Levine's own analysis, couched in a variant of type-logical categorial grammar, relies on a gapping-specific conjunction to account for the surface asymmetry between the gapped and source conjuncts. But positing such gapping-specific conjunction makes it difficult to capture the generalization that gapping is possible with a variety of connectives, including certain subordinators that convey semantic parallelism (Culicover & Jackendoff 2005).

(where the negated modal contributed by *can't* takes wide-scope in the interpretation). But this is unexpected if we take conjuncts in gapping to be sometimes a vP.

(9) *John can't eat a Big Mac and merely Mary a spoonful of soup.

Thus, prior syntactic analyses have failed to produce an adequate analysis of gapping. On the one hand, there is evidence against the supposed vP coordination structure. On the other hand, the connection purported to exist between the site of coordination and the availability of one reading or another does not always hold. The empirical evidence we have discussed so far does not seem to support a treatment of gapping as a syntactic ambiguity phenomenon.

3 Scope ambiguity in gapping is a semantic ambiguity

I propose an alternative, semantic analysis for the distinction between wide- and distributive-scope readings of gapping sentences. In my analysis, the two readings of 1 result from different ways of resolving a single, underspecified meaning (Copestake *et al.* 2005; Egg *et al.* 2001; Reyle 1993; Richter & Sailer 2004). Since this underspecified meaning can be linked to a single, uniform syntactic structure, ambiguous readings of gapping sentences do not need to rely on a syntactic ambiguity. This is a major advantage over previous syntactic analyses discussed above, where scopally ambiguous readings of a sentence necessarily require multiple distinct syntactic structures assigned to that sentence.

I assume that gapped conjuncts are clausal constituents (invariably of category S) and that they are underspecified for their semantic type. I also assume that the semantics of coordination is underspecified in that what is being conjoined between the first and second conjuncts can be a subterm of the first conjunct, provided that the conjoined terms match in their semantic type. The two readings of gapping sentences will then arise as the result of resolving these underspecified meanings.

More specifically, I assume a type-logical distinction between tensed and untensed clauses that is more-or-less standard. The particular system I adopt in this paper is one proposed in Champollion 2015, although nothing crucial in my analysis hinges on this choice. In Champollion's semantics, tensed clauses are of type t and tenseless clauses are of type $\langle vt, t\rangle$ (properties of sets of events); operators such as negation and modals are of type $\langle\langle vt, t\rangle, \langle vt, t\rangle\rangle$ (modifiers of event descriptions). From this the following two possibilities are predicted: A wide-scope reading (such as 2a) results when the conjoined terms are of type $\langle vt, t\rangle$, and the tense semantics as well as scopal operators in the first conjunct outscope the conjunction; a distributive-scope reading (such as 2b) results when the conjoined terms are of type t, and each conjunct includes its own tense and scopal operators.

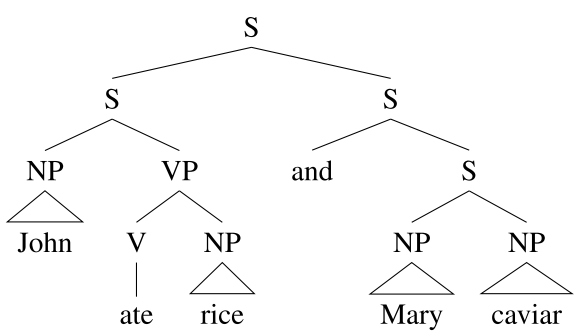
My proposal is compatible with a number of different hypotheses about the syntax of gapped clauses and the mechanism that derives the interpretation at the gap. Without going further into these issues, I will provide just one way in which my proposal can be implemented. First, following Culicover & Jackendoff (2005) and Abeillé *et al.* (2014), I analyze gapped clauses as sentence fragments licensed by the following rule (where $^+$ represents a Kleene plus).

(10) $\mathrm{S} \rightarrow \mathrm{XP}\ \ \mathrm{XP}^+$ — where each XP is a focus

This rule licenses non-headed clauses that contain two or more focal daughters. When these fragment clauses are embedded in a non-initial conjunct, the result is gapping.

(11) a. John ate rice and Mary caviar.

b.



Next, turning to the semantics of gapping, I adopt the approach to ellipsis proposed in Ginzburg & Sag 2000 and assume that the content of a gapped clause is obtained by combining the content of its daughters and an (implicit) question under discussion (QUD) which, in a sense, the gapped and source clauses provide an answer to.[6] The interpretation of an entire gapping sentence will follow from independently motivated rules of the syntax and semantics of clause conjoining and the way the content of gapped clauses is determined, as will be seen in 4.2.

4 A semantic underspecification-based analysis of gapping

In this section, I present an analysis of gapping in an underspecified semantic representation and show how the wide- and distributive-scope readings are licensed. I first introduce the semantic framework in 4.1. An example analysis of gapping in conjunctive coordination is provided in 4.2.

4.1 Semantic underspecification and Lexical Resource Semantics

While classical formal semantics uses fully specified logical formulas, underspecified semantics uses meta-representations of partially specified formulas which subsume a set of resolved logical formulas. As is well known, the presence of quantifiers triggers multiple different readings. For example, the sentence in 12 has two different readings, one in which each student read a different book (∀>∃) and a second reading in which there is a single book and all students read it (∃>∀). These readings are represented in 12a-b.

(12) Every student read a book.

a. $\forall x[\text{student}(x) \rightarrow \exists y[\text{book}(y) \wedge \text{read}(x, y)]]$

b. $\exists y[\text{book}(y) \wedge \forall x[\text{student}(x) \rightarrow \text{read}(x, y)]]$

These representations differ only in the permutations of the quantifiers. Semantic underspecification abstracts away from these differences by providing a unique meta-representation in which the differences are omitted. An underspecified description of 12

[6] I leave open the possibility of a unified QUD-based analysis for gapping and so-called pair-list answers, such as (i).

(i) A: Who ate what? B: John, the Big Mac; Mary, just the fries.

For motivations and advantages for such unified analysis, see Reich 2007.

is exemplified in Figure 1, where the relative scopal relations between the quantifiers is unspecified.

$$\boxed{1}\,\forall x(\text{student}(x) \to \boxed{2}\,) \qquad\qquad \boxed{3}\,\exists \text{y}(\text{book}(y) \wedge \boxed{4}\,)$$

$$\boxed{5}\,\text{read}(x, y)$$

Figure 1: Underspecified description of 12

The arrows in Figure 1 indicate subordination constraints. Thus, this description states that no matter how the scope is resolved, the verbal nucleus ($\boxed{5}$) must be within the scope of both quantifiers. The description allows for two solutions: 12a is obtained if $\boxed{2}$ is equated with $\boxed{3}$, and 12b is obtained if $\boxed{4}$ is equated with $\boxed{1}$.

There are several different underspecified semantic formalisms available, but in this work I adopt Lexical Resource Semantics (LRS, Richter & Sailer 2004). The core idea of LRS is that the semantic representations of sentences result from collecting the meaning contributions of words in accordance with combinatorial principles that constrain semantic composition.[7] LRS distinguishes among different aspects of the semantic representation associated with a sign: EX(TERNAL-)CONT(ENT) indicates the overall logical form of a sign; IN(TERNAL-)CONT(ENT) encodes that part of the logical form of a sign outscoped by any other operator within the syntactic projection of that sign; PARTS is a collection of the meaning contributions of words.

To show how semantic composition is carried out in LRS, a partial LRS analysis of 12 is provided in Figure 2 (see next page). The particular combinatorial principles that govern the combination of quantifiers with their heads are stated in 13.

(13) The LRS Semantics Principle

1. If the nonhead is a quantifying determiner, then the INCONT value of the head is a subterm of the restrictor.

[7]Unlike most other semantic formalisms, LRS does not employ the lambda calculus as its compositional mechanism.

2. If the nonhead is a quantified NP, then the INCONT value of the head is a subterm of the scope.

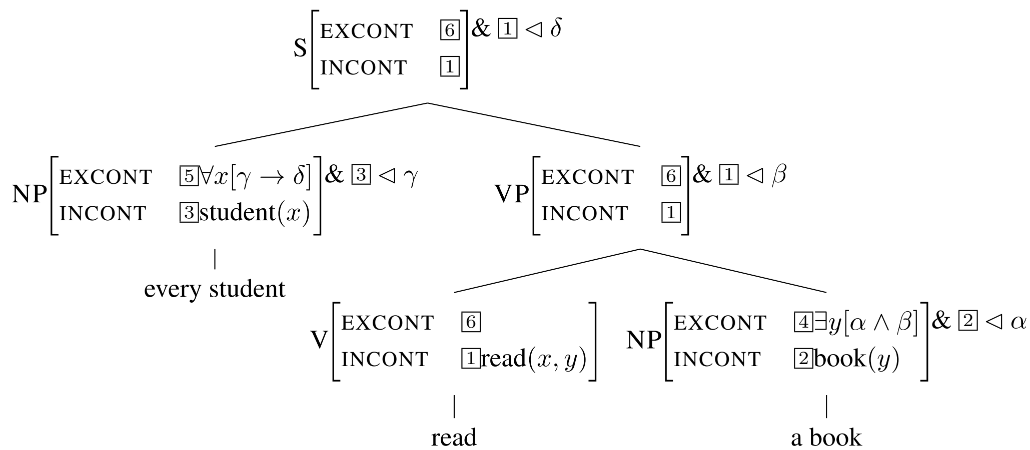


Figure 2: An LRS analysis of 12

In this figure, the restrictor and scope of quantifiers are not known and are given in terms of meta-variables (represented with Greek letters). By courtesy of the first clause of the LRS Semantics Principle, the INCONT of the head nouns *student* and *book* is required to be a subterm of the restrictor ($\boxed{3} \lhd \gamma$ and $\boxed{2} \lhd \alpha$, where $\lhd$ means 'is a subterm of'). The combination of the quantified NPs and their respective heads is constrained by the second clause of the LRS Semantics Principle ($\boxed{1} \lhd \delta$ and $\boxed{1} \lhd \beta$); this ensures that the semantic contribution of the verb is contained in the scope of the quantifiers.

Turning to the verb projections, the verb *read* imposes no restriction on the EXCONT at the lexical level ($\boxed{6}$), and the EXCONT and INCONT of the verb are identified along the verb's syntactic projections (due to the LRS Projection Principle[8]). Because the EXCONT of a maximal projection must be constructed from all semantic contributions of all the words in it (due to the LRS Excont Principle[9]), there are exactly two possible solutions for $\boxed{6}$: $\boxed{6}$ can be equated with $\boxed{5}$ (and the $\forall>\exists$ reading results), or it can be equated with $\boxed{4}$ (and the $\exists>\forall$ reading results).

[8]Richter & Kallmeyer 2009:47–48.
[9]Richter & Kallmeyer 2009:47.

4.2 Gapping in conjunctive coordination

Recall that the crucial assumption that accounts for the scope ambiguity in gapping is the asymmetry of semantics between the first and second conjuncts: An operator in the first conjunct can possibly outscope the conjunction, but one in the second conjunct cannot. This asymmetry is not restricted to gapping.[10]

(14) a. There is **no** medicine or any treatment whatsoever.

b. *There is any medicine or **no** treatment whatsoever.

In 14a the NPI *any* in the second conjunct is licensed by the negation *no* in the first conjunct. But a negation in the second conjunct does not seem to be able to license an NPI in the first conjunct, as can be seen from 14b. This semantic asymmetry between the first and second conjuncts will follow from a general LRS constraint on coordination, which is formulated in 15.

(15) The LRS Semantics Principle (Clause 3.)
In coordinate structure, the INCONT of the first conjunct must be a subterm of the first argument of conjunction (α in $\alpha \wedge \beta$), and the EXCONT of the second conjunct must be a subterm of the second argument of conjunction (β).

As a result of this constraint, when two elements, say A and B, are conjoined (to form 'A *and B*'), every subterm of B will be contained in the interpretation of the conjunction, whereas a scopal element within A can possibly outscope the conjunction. This allows for the following two possibilities: Either the conjunction takes widest scope (and each conjunct is interpreted as scopally independent), or a scopal element within the first conjunct takes widest scope.

With this constraint on coordination and the type-logical differences between tensed and untensed clauses (Champollion 2015) in place, we are ready to account for the scope ambiguity in gapping. In Figure 3 I provide a partial LRS analysis of 1 *Max can't live in San Francisco and Claire in Buffalo*.[11]

[10] I thank Jean-Pierre Koenig for providing these examples.

[11] The analysis of clause-internal syntax and semantics is omitted here, as it does not bear on the issue of the scope ambiguity in gapping and rather follows trivially from the LRS syntax-semantics interface.

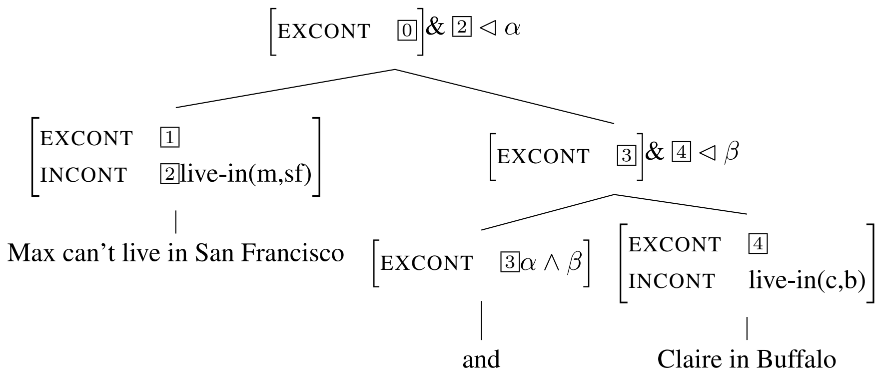


Figure 3: An LRS analysis of 1

In this figure, the EXCONT values of both conjuncts as well as that of the topmost mother node are not known. But whichever way these values are determined, the EXCONT of the second conjunct must be included in the second argument of the conjunction ($\boxed{4} \lhd \beta$), while for the first conjunct, we only require that its INCONT be included in the first argument of the conjunction ($\boxed{2} \lhd \alpha$). Moreover, because the EXCONT of the top mother must be constructed from all semantic contribution of all of its daughters, the EXCONT of the top mother must include the semantic contribution of *can't*, which includes tense as well as the negative modal. Since modals combine with untensed clauses and occur in the scope of tense (Champollion 2015; Condoravdi 2002), the following two possibilities are predicted: The wide-scope reading of 1 results when the EXCONT value of the first conjunct is identified with the EXCONT value of the entire sentence($\boxed{1} \equiv \boxed{0}$); and the distributive-scope reading results when the conjunction meaning is identified with the EXCONT value of the entire sentence ($\boxed{3} \equiv \boxed{0}$). These readings are represented in 16a-b.

(16) a. TENSE[$\neg\Diamond$live-in(m, *sf*) $\wedge$ $\neg\Diamond$live-in(c, b)] ($\boxed{1} \equiv \boxed{0}$)

b. TENSE[$\neg\Diamond$live-in(m, *sf*)] $\wedge$ TENSE[$\neg\Diamond$live-in(c, b)] ($\boxed{3} \equiv \boxed{0}$)

Finally, the fact that there is no wide-scope reading for a sentence such as 3 *Max can't live in San Francisco and Claire in Buffalo* follows from the constraint that the semantic type of conjoined terms must match (Partee & Rooth 1983) and the fact that tensed clauses are of type t.

(17) ill-formed semantic representation of 3:
$*\textsc{tense}[\neg\Diamond\text{live-in}(m, \textit{sf}) \wedge \textsc{tense}(\neg\Diamond\text{live-in}(c, b))]$

5 Conclusion

Scope ambiguities in gapping have long presented a challenge to theories of the syntax-semantics interface, as scopal elements do not appear where they are interpreted. In this paper, I have expanded on the use of semantic underspecification to model scope ambiguities in gapping (Copestake *et al.* 2005; Egg *et al.* 2001; Reyle 1993; Richter & Sailer 2004) and proposed that these ambiguities result from the underspecification of what is being conjoined and independently motivated semantic types for clauses (Champollion 2015). The analysis presented in this paper does not rely on a syntactic ambiguity (such as in Potter *et al.* 2017) or any gapping-specific assumption to account for the various readings of gapping sentences, and it therefore allows a simpler account of the interaction between gapping and scope.

REFERENCES

ABEILLÉ, ANNE, BÎLBÎIE, GABRIELA, & MOURET, FRANÇOIS. 2014. A Romance perspective on gapping constructions. *Pages 227–267 of:* BOAS, H., & F., GONZÁLVEZ-GARCÍA (eds), *Romance in Construction Grammar*. John Benjamins.

CHAMPOLLION, LUCAS. 2015. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy*, **38**, 31–66.

CONDORAVDI, CLEO. 2002. Temporal interpretation of modals: Modals for the present and for the past. *Pages 59–87 of:* BEAVER, D., MARTÍNEZ, L. C., CLARK, B., & KAUFMANN, S. (eds), *The Construction of Meaning*. CSLI Publications. Stanford, CA.

COPESTAKE, ANN, FLICKINGER, DAN, POLLARD, CARL, & SAG, IVAN A. 2005. Minimal recursion semantics: An introduction. *Pages 281–332 of: Research on Language and Computation*, vol. 3. Kluwer Academic Publishers.

COPPOCK, ELIZABETH. 2001. Gapping: In defense of deletion. *Chicago Linguistics Society*, **37**, 133–148.

CULICOVER, PETER W., & JACKENDOFF, RAY. 2005. *Simpler Syntax*. Oxford, New York: Oxford University Press.

EGG, MARKUS, KOLLER, ALEXANDER, & NIEHREN, JOACHIM. 2001. The constraint language for lambda structures. *Journal of Logic, Language and Information*, **10**, 457–485.

GINZBURG, JONATHAN, & SAG, IVAN A. 2000. *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. CSLI Publications.

JOHNSON, KYLE. 2009. Gapping is not (VP-) ellipsis. *Linguistic Inquiry*, **40**, 289–328.

KUBOTA, YUSUKE, & LEVINE, ROBERT. 2016. Gapping as hypothetical reasoning. *Natural Language and Linguistic Theory*, **34**(1), 107–156.

LIN, VIVIAN I-WEN. 2002. *Coordination and sharing at the interfaces*. Ph.D. thesis, MIT.

LÓPEZ, LUIS, & WINKLER, SUSANNE. 2003. Variation at the syntax-semantics interface: Evidence from Gapping. *Pages 227–248 of:* SCHWABE, K., & WINKLER, S. (eds), *The Interfaces: Deriving and interpreting omitted structures*. Linguistik Aktuell 61. Amsterdam: John Benjamins Publishing Company.

PARTEE, BARBARA H., & ROOTH, MATS. 1983. Generalized conjunction and type ambiguity. *Pages 361–383 of:* BAUERLE, R., SCHWARTZ, C., & VON STECHOW, A. (eds), *Meaning, Use, and Interpretation of language*. Berlin: De Gruyter.

PESETSKY, DAVID. 1997. Optimality theory and syntax: Movement and pronunciation. *Pages 134–170 of:* ARCHANGELI, D., & LANGENDOEN, D. TERENCE (eds), *Optimality Theory: An overview*. Oxford: Blackwell.

POTTER, DAVID. 2014. A heterogeneous approach to gapping. *In: Proceedings of the 31st West Coast Conference on Formal Linguistics*.

POTTER, DAVID, FRAZIER, MICHAEL, & YOSHIDA, MASAYA. 2017. A two-source hypothesis for Gapping. *Natural Language and Linguistic Theory*, **35**(4), 1123–1160.

REICH, INGO. 2007. Toward a uniform analysis of short answers and gapping. *Pages 467–84 of:* SCHWABE, K., & WINKLER, S. (eds), *On Information Structure, Meaning and Form*. Amsterdam: John Benjamins.

REPP, SOPHIE. 2009. *Negation in Gapping*. Oxford Studies in Theoretical Linguistics, vol. 22. Oxford University Press.

REYLE, UWE. 1993. Dealing with ambiguities by underspecification: Construction, representation and deduction. *Journal of Semantics*, **10**, 123–179.

RICHTER, FRANK, & KALLMEYER, LAURA. 2009. Feature logic-based semantic composition: A comparison between LRS and LTAG. *Pages 32–92 of: Typed feature structure grammars*. Frankfurt a.M.: Peter Lang.

RICHTER, FRANK, & SAILER, MANFRED. 2004. Basic concepts of lexical resource semantics. *Pages 87–143 of: European Summer School in Logic, Language and Information 2003. Course Material I, Volume 5 of Collegium Logicum*. Kurt Gödel Society Wien.

ROCHEMONT, MICHAEL, & CULICOVER, PETER. 1990. *English Focus Constructions and the Theory of Grammar*. Cambridge: Cambridge University Press.

SAG, IVAN. 1976. *Deletion and Logical Form*. Ph.D. thesis, Massachusetts Institute of Technology.

SIEGEL, MUFFY E. A. 1984. Gapping and interpretation. *Linguistic Inquiry*, **15**, 523–530.

TOOSARVANDANI, MAZIAR. 2013. Gapping is low coordination (plus VP-ellipsis): A reply to Johnson. Ms., MIT.

Phonological environment and the social evaluation of American English sibilants

JACOB B. PHILLIPS
Phonology Laboratory, University of Chicago[*]

1 Introduction

Many speakers of English exhibit /s/-retraction, the process by which /s/ is produced or perceived approaching /ʃ/ in the context of /ɹ/, especially in /stɹ/ clusters. Thus for a retractor, the word *street* is pronounced approaching *shtreet*. This has been robustly observed across different dialect groups of American English, including New York/Long Island (Kraljic 2008), Philadelphia (Gylfadottir 2015), Raleigh, North Carolina (Wilbanks 2017), Columbus, Ohio (Durian 2007), and southern Louisiana (Rutter 2011). It is so widespread that it has been referred to as a 'general American innovation' (Shapiro 1995), although it has additionally been observed outside the U.S., including Newfoundland, Canada (Clarke 2004). New Zealand (Lawrence 2000), England (Bass 2009, Glain 2014) and Scotland (Stuart-Smith et al. 2018). Unlike typical sound changes, there does not appear to be a gender effect in /s/-retraction production (Gylfadottir 2015), although women were found to lead the change in Raleigh (Wilbanks 2017).

Furthermore, while retraction is observed in /stɹ/ clusters across these different dialects of English, the same retraction is not observed in /spɹ/ or /skɹ/ clusters, either in frequency or rate of retraction, such that *scream* is rarely pronounced approaching *shcream* and *spree* is rarely pronounced approaching *shpree* (Shapiro 1995, Baker et al. 2011, Gylfadottir 2015). This asymmetric distribution poses a challenge to many accounts of the sound change, such as those that rely on overlapping gestures between the /s/ and the /ɹ/, predicting retraction in /spɹ/ clusters as the intervening /p/ has no associated lingual gesture. The observed asymmetry has received some attention examining the phonetic nature of the dis-

[*]This work would not have been possible without the contributions and invaluable research assistance of Hillel Steinmetz. I would like to thank Alan C. L. Yu, Laura Staum Casasanto and audiences at NWAV 46 for their advice, feedback and encouragement. I would also like to thank Alex Kramer for technical support and Giovanna Hooton and Paige Resnick for their help running participants in the laboratory. Funding in part provided by the University of Chicago Graduate Aid in Linguistics fund and the Sapir Scholarship program at the University of Chicago.

tribution, for example proposing that the affrication of /t/ is the driving force of retraction (Lawrence 2000), and similarly in phonological accounts, proposing that the underspecification of a place feature of /t/ allows for the assimilation across /t/ but not /{p, k}/ (Shapiro 1995). Yet little work has sough to examine the socioindexical nature of this variable and whether that index varies by phonological context.

While most of the studies examining /s/-retraction have focused primarily on production data, a small but growing body of work has looked at the perception and social evaluation of the phenomenon. For example, listeners have been shown to not shift their /s/-/ʃ/ category boundaries after exposure to retraction in /stɹ/ clusters, like *district*, but do shift those boundaries if retraction is observed in unexpected environments, like *colosseum* (Kraljic et al. 2008), suggesting an attention to ongoing sound changing and/or compensation for coarticulation. Recent research in Austin, Texas found that /s/-retraction is below the level of speakers' consciousness, with most speakers unaware of the phenomenon when questioned in interviews (Hinrichs et al. 2015). When prompted to give social meaning to the phenomenon, speakers provided diverse responses, often suggesting that it may be more Southern/hick, non-native, or the result of speech impediment. However, no studies to date have examined listener's implicit social evaluation of /s/-retraction, either in /stɹ/ or /s{p, k}ɹ/ clusters.

In contrast, much work in sociolinguistics has rigorously examined the social evaluation of predominantly prevocalic sibilant variation in English. This indexical relationship has been found on various dimensions, with a fronted /s/ typically perceived as more middle class, feminine (Stuart-Smith 2007) and gay (Levon 2014). Similarly, a retracted /s/ has been demonstrated to be perceived as more rural/country (Campbell-Kibler 2011, Podseva & Van Hofwegen 2014), masculine (Zimman 2013) and lesbian (Podseva & Van Hofwegen 2014). Additionally the socioindexical meaning of /s/ variation has been found to vary by social context, indexing sexuality and gender in some contexts but not others (Pharao 2014). However, as these studies of sibilant variation generally focus on prevocalic /s/, they are largely agnostic to any role that the phonological environment may play in determining the socioindexical meaning of /s/.

The present study consists of a social evaluation ratings task that seeks to examine the socioindexical meaning of /s/ in /s{p, t, k}ɹ/ clusters in order to better understand the complicated nature of the sound change. Specifically, this study asks if /s/ in /stɹ/ clusters contributes different socioindexical meaning than /s/ in /s{p, k}ɹ/ clusters that can provide a better understanding of the asymmetrical phonological distribution as well as the lack of an expected gender effect in many communities.

2 Methods & materials

Stimuli: Eight model talkers (4 male, mean age 20) were recruited to record the auditory stimuli. All model talkers were native speakers of American English (four California, four Illinois) and received either payment or credit for an introductory linguistics course. The model talkers were seated in an isolated double-walled sound booth and were recorded on a Zoom H4n recorder with a Shure SM10A head-mounted microphone as they read a series of stop- and sibilant-initial lexical words off the computer screen in the carrier phrase: *Please say string again. String string string.* The wordlist for the sibilant-initial stimuli is provided in Table 1. The eight model talkers showed varying degrees of retraction in /stɹ/ clusters, with one male and one female exhibiting forms categorically perceived by the researchers to be /ʃ/ rather than /s/, i.e. the retractors.

Table 1: Wordlist for stimuli creation

	/s/	/sC/	/sCɹ/	/ʃ/
/p/	sit	spit	spritz	shit
/t/	sing	sting	string	shin
/k/	sip	skip	script	ship

To create the /s{p t k}ɹ/ stimuli for the ratings task, the sibilant onsets from the corresponding minimal pairs were mixed digitally mixed using a Praat script (Darwin 2005) at four scaling ratios: 0%-/s/:100%-/ʃ/, 60%-/s/:40%-/ʃ/, 80%-/s/:20%-/ʃ/, and 100%-/s/:0%-/ʃ/. The scaling ratios were selected to include the extreme endpoints (0% and 100% /s/) and more naturalistic retraction values (60% and 80% /s/). Thus, for the word *script* the initial /s/ was removed and cross-spliced with the an onset mixed from *sip* and *ship*. Separate onset continua were then created for *spritz* and *string*. Stop-initial stimuli were included as fillers but were not manipulated. All files were normalized for intensity.

Participants: 342 (208 males, mean age 30) geographically diverse (113 urban, 179 suburban, 50 rural) native speakers of American English were recruited on Mechanical Turk and were paid for their participation in this study. Gender, age, and geographic environment were self-reported by the participants. No participants reported speech or hearing disorders/abnormalities. 41 additional participants were recruited but excluded from the analysis due to technical difficulties or non-attentive responses.

Procedure: In order to identify the attributes to include in the social evaluation ratings task, an initial open-ended panel study was conducted following on Amazon Mechanical Turk with 15 (10 male) native speakers of American English. Panel participants heard the unmanipulated sibilant-initial stimuli presented in Table 1 for each model talker and were asked to supply any information about the speaker's identity, characteristics or traits in an open-ended question. For the male retractor, panel participants were split in describing his /stɹ/ tokens as more masculine, athletic and straight than his other /s/ tokens or whether those same tokens were more gay, intellectual, and pretentious. The female retractor did not elicit strong changes in panel participants' responses in her /stɹ/ clusters, which were typically described as younger. Following the panel study seven attributes were selected including attractiveness, masculinity, friendliness, formality, shyness and sexuality. Additionally, geographic regions and environment types were included.

Table 2: Sample survey matrix

Please provide your impression of the speaker on the scales below:

Attractive	○	○	○	○	○	○	○	Unattractive
Masculine	○	○	○	○	○	○	○	Not at all masculine
Unfriendly	○	○	○	○	○	○	○	Friendly
Casual	○	○	○	○	○	○	○	Formal
Shy	○	○	○	○	○	○	○	Outgoing
Heterosexual	○	○	○	○	○	○	○	Homosexual

Where does it sound like the speaker might be from? (select all that apply):

☐ The South	☐ The Midwest	☐ The Northeast
☐ West coast	☐ East coast	☐ The Southwest
☐ The City	☐ The Country	☐ The Suburbs

In the social evaluation ratings task, participants were assigned to one of the four retraction conditions (0%, 60%, 80% or 100% /s/) for one of the three clusters (/spɹ/, /stɹ/ or /skɹ/), in order to allow for the different clusters to be rated and judged independently. Participants heard a total of 24 audio clips, with 8 targets and 16 fillers. Participants completed the task online and were instructed to use headphones. Each audio clip played once automatically and participants could optionally replay the clip as many times as they chose. After listening to each clip, participants were asked to rate the talker on each of the selected attributes presented on seven-point scales in. The order and polarity of the attributes were randomized across participants, but held constant between trials for the same participant. Participants were additionally asked to select radio buttons corresponding to geographic categories. The order of the radio buttons was constant across participants and trials. A sample survey matrix is provided in Table 2.

Analysis: Linear mixed effects regression models were fitted separately for the different attributes (z-scored by participant), using the `lmer()` function of the `lme4` package in R, but only results for SEXUALITY and MASCULINITY will be presented here as those responses were most strongly conditioned by /s/ variation. Separate models were run for the male and female model talkers to allow for gender-specific differences in the evaluation of the different attributes. The models were constructed using a nested approach, incrementally adding variables as either fixed or random effects. Variables that improved model likelihood in any of the models were included in the models for the other attributes or gender. The fixed effects included in the analysis were WORD (SPRITZ, STRING, SCRIPT; treatment coded with STRING as base), retraction CONDITION (0, 60, 80, 100; scaled) and SPEAKER (F1, F2, F3, F4, M1, M2, M3, M4; treatment coded with F1 and M1 as base). Two-way interactions were included between the fixed effects. Random intercepts were included for TRIAL (scaled) and LISTENER. The state of origin of the model talkers as well as self-reported demographic categories of the listener (e.g. age, gender, geographic information) did not significantly improve model likelihood. To further examine the role of individual variability captured by SPEAKER, separate linear regressions were on each model talker with WORD and CONDITION as independent variables.

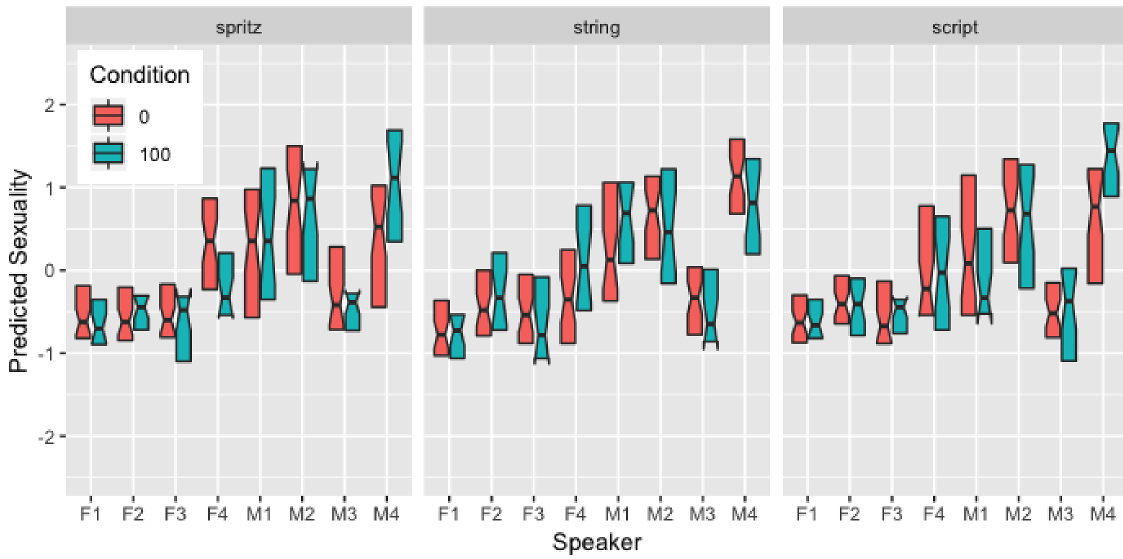


Figure 1: Predicted sexuality (z-scored, y-axis) by model talker (x-axis), condition (color) and word (panel).A positive predicted sexuality value indicates a more HOMOSEXUAL response, while a negative value indicated a more HETEROSEXUAL response.

3 Results

Sexuality: For male model talkers, there was no main effect of WORD, CONDITION or the interaction of the two variables at the community level. There was a significant effect of SPEAKER, with model talker M3 more likely to be rated by listeners as more heterosexual ($\beta = -0.53, p < 0.001$) and model talker M4 as more homosexual ($\beta = 0.26, p < 0.05$).

For female model talkers, like male model talkers, there was no significant effect of WORD or CONDITION. There was a significant effect of SPEAKER, with model talkers F2 and F4 more likely to be evaluated as more homosexual (for F2: $\beta = 0.63, p < 0.05$; for F4: $\beta = 0.98, p < 0.01$). Figure 1 provides the predicted SEXUALITY ratings for each of the model talkers in the endpoint conditions (0 and 100% /s/), with listeners exhibiting significant influences of CONDITION or WORD in the evaluation of model talkers F4 and M4, both more likely to be rated as more homosexual, but not for the other model talkers.

Model talker M4, a male from Illinois, was significantly more likely to receive more homosexual ratings than the other males in the study. Additionally, M4 alone shows any

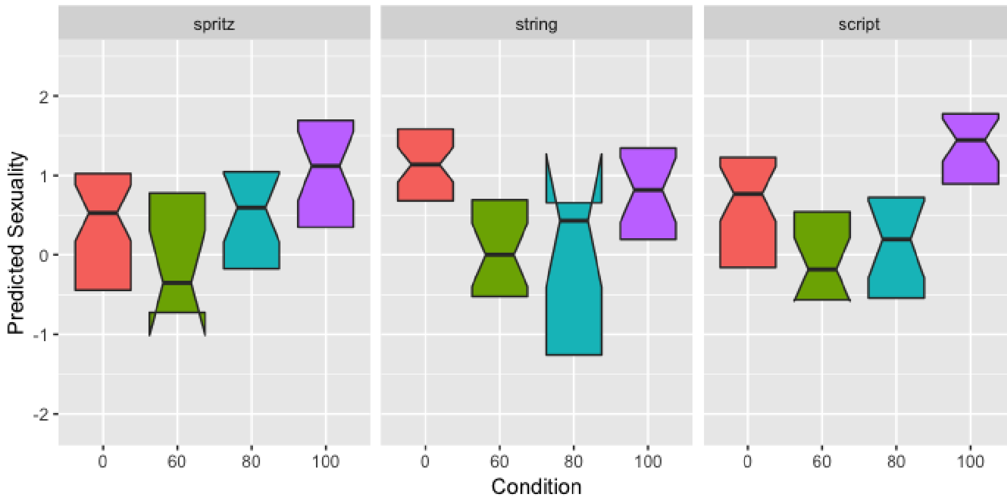


Figure 2: Predicted sexuality (z-scored, y-axis) of model talker M4, by condition (x-axis,color) and word (panel). A positive predicted sexuality value indicates a more HOMOSEXUAL response, while a negative value indicated a more HETEROSEXUAL response.

effect of WORD and CONDITION in the individual linear regression. For M4, there is a significant interaction of WORD and CONDITION with listeners significantly more likely to give a more heterosexual response in increased retraction conditions in the /spɹ/ and /skɹ/ clusters than in the /stɹ/ clusters (for /p/: $\beta = -0.39, p < 0.001$; for /k/: $\beta = -0.36, p < 0.01$). Figure 2 illustrates listener's predicted sexuality responses for M4 in all conditions.

Model talker F4, a female from Illinois, was also significantly more likely receive more homosexual ratings than other female model talkers in the study. Like M4, F4 was the lone female talker to show significant effects of WORD and CONDITION in the individual linear regression. The observed main effect of CONDITION suggests that F4 was more likely to receive to more heterosexual responses in increased retraction conditions ($\beta = -0.10, p < 0.05$). The additional interaction of WORD and CONDITION counteracts the main effect of CONDITION, suggesting that she received more homosexual ratings in retracted conditions for /spɹ/ clusters compared to /stɹ/ clusters ($\beta = 0.32, p < 0.05$). Although it was trending in the same direction for /skɹ/ clusters, it was not significant ($\beta = 0.20, p = 0.13$). Figure 3 illustrates listener's predicted sexuality responses for F4 in all conditions.

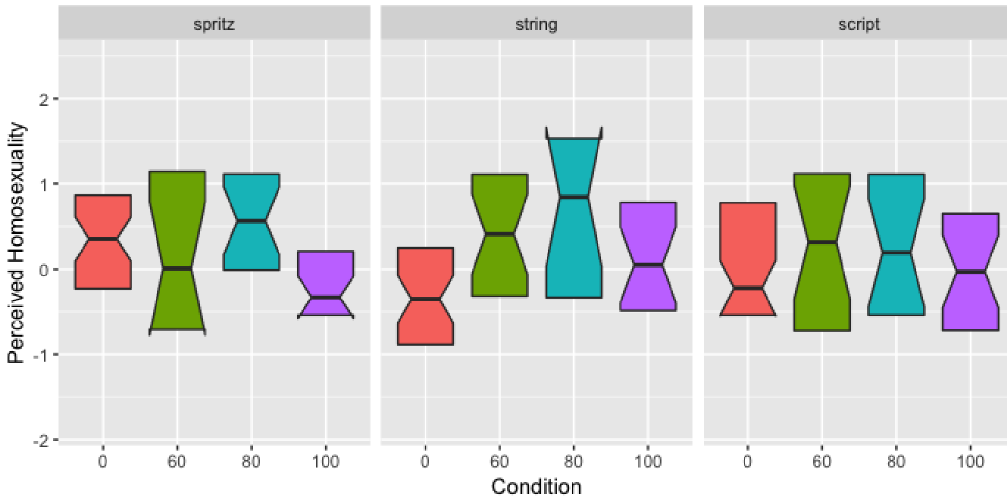


Figure 3: Predicted sexuality (z-scored, y-axis) of model talker F4, by condition (x-axis,color) and word (panel). A positive predicted sexuality value indicates a more HOMOSEXUAL response, while a negative value indicated a more HETEROSEXUAL response.

Masculinity: For male model talkers, there was an observed main effect of SPEAKER, with model talkers M2 and M4 rated as less masculine (for M2: $\beta = -0.26, p < 0.05$; for M4: $\beta = -0.44, p < 0.001$). There was additionally a main effect of WORD with /skɹ/ clusters being rated by listeners as less masculine than /stɹ/ clusters ($\beta = -0.26, p < 0.05$). The interaction of WORD and CONDITION suggests that listeners are significantly less likely to give more masculine responses to /skɹ/ clusters in increased retraction conditions than /stɹ/ clusters ($\beta = -0.13, p < 0.05$). However, these effects were counteracts by the interaction of SPEAKER and WORD, with speaker M3 more likely to receive more masculine responses in /skɹ/ words ($\beta = 0.33, p < 0.05$).

For female talkers, there was again a significant effect of SPEAKER, with model talkers F2 and F4 more likely to be evaluated as more masculine (for F2: $\beta = 0.21, p < 0.05$; for F4: $\beta = 0.50, p < 0.001$). Like for the male talkers, there was also a significant interaction of WORD and CONDITION, suggesting that listeners are significantly more likely to give more masculine responses to /skɹ/ clusters in increased retraction conditions than /stɹ/ clusters, in contrast to the effect for male talkers ($\beta = 0.13, p < 0.05$).

Model talker M1, a male from California, was consistently rated as above average for

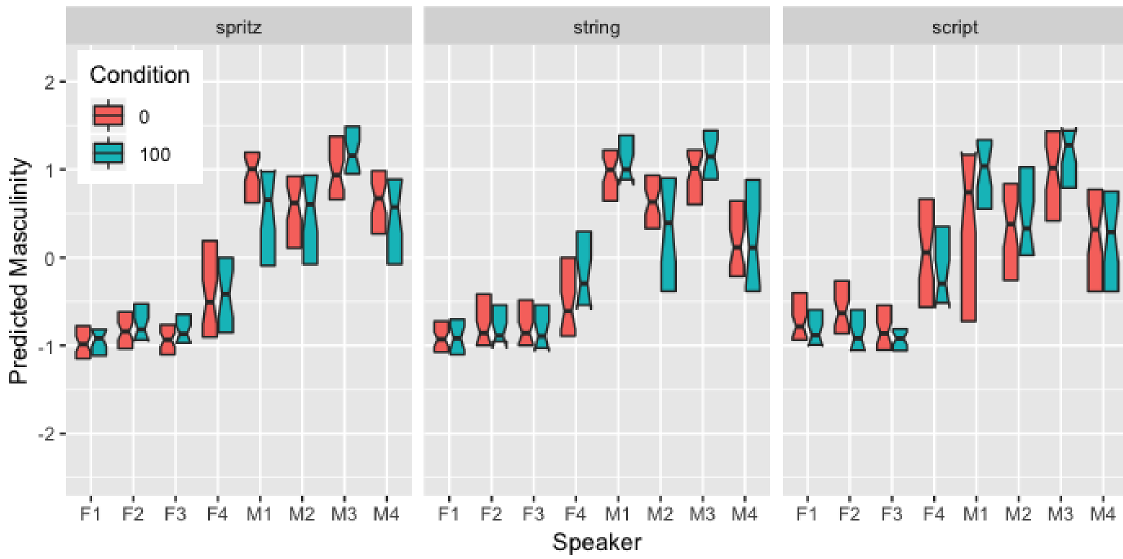


Figure 4: Predicted masculinity (z-scored, y-axis) by model talker (x-axis), condition (color) and word (panel). A positive predicted sexuality value indicates a more MASCULINE response, while a negative value indicated a less MASCULINE response.

masculinity compared to the other model talkers. M1 exhibited a main effect of WORD, with listeners significantly more like to rate his /skɹ/ clusters as less masculine than his /stɹ/ clusters ($\beta = -0.27, p < 0.01$). This effect was further amplified by the interaction of WORD and CONDITION, as listeners were more likely to rate his /skɹ/ clusters as less masculine in increased retraction conditions than his /stɹ/ clusters ($\beta = -0.21, p < 0.05$).

Again, model talker F4, significantly more likely to be rated as more masculine than the other female model talkers, stood out in the individual linear regressions as the only female model talker to exhibit significant effects of WORD or CONDITION. The main effect of WORD suggests that listeners perceived /skɹ/ clusters produced by F4 to be more masculine than her /stɹ/ clusters ($\beta = 0.32, p < 0.01$). This effect is reinforced by the interaction of WORD and CONDITION, as F4's /skɹ/ clusters are significantly more likely to be rated as more masculine in increased retraction conditions as compared to her /stɹ/ clusters ($\beta = 0.21, p < 0.05$).

4 Discussion

The goals of this study were to examine the ways in which phonological environment of /s/, that is the different clusters /s{p, t, k}ɹ/, contribute to the social meaning of /s/-retraction, and conversely, the ways in which social evaluation of /s/ in these clusters can shed light on the asymmetric distribution of /s/-retraction. The results of the ratings task suggest that phonological environment plays a nuanced role in how listeners evaluate the sexuality and masculinity of different speakers.

For the evaluation of perceived sexuality, not only do individual participants vary significantly in how they rate the model talkers, but also individual speakers vary consistently in how they are rated. For speakers who are less likely to be perceived as homosexual, the phonological environment and the degree of retraction of /s/ do not further condition the evaluation as more or less homosexual. However, for speakers more likely to be perceived as gay or lesbian, the interaction of the environment and the degree of retraction significantly influence listeners' perceptions of the speaker's sexuality. For male speakers (specifically M4), they are more likely to be perceived as more straight in /spɹ/ and /skɹ/ clusters in increased retraction conditions. This suggests that a retracted /s/ in these clusters has the expected socioindexical meaning that a retracted /s/ carries in prevocalic environment, that is, more straight (for men). However, in /stɹ/ clusters, increased retraction does not contribute to perceived straightness. The same pattern, albeit in reverse, is observed in female speakers more likely to be perceived as lesbian (F4): Increased retraction in /spɹ/ clusters contributes to a stronger perceived homosexuality than in /stɹ/ clusters. Taken together, these results suggest that a retracted /s/ specifically in /stɹ/ clusters does not index straightness (for men) or gayness (for women).

For the evaluation of perceived masculinity, the results for female talkers align with the perceived sexuality results. That is, a retracted /s/ in /skɹ/ clusters is perceived as more masculine than a retracted /s/ in /stɹ/ clusters. This effect emerged individually as well for F4, a woman perceived as more lesbian and more masculine than her peers. However, the results for male speakers illustrate the reverse pattern, with a retracted /s/ in /skɹ/ clusters being evaluated as less masculine than a retracted /s/ in /stɹ/ clusters,. This pattern also emerged individually for M1, a male perceived as more masculine than his peers. Taken together these results can be interpreted in a few different ways. Firstly, it may be the case that a retracted /s/ in /skɹ/ clusters indexes gender atypicality. Or secondly, it may be that the dampening of socioindexical value of /stɹ/ is of more consequence in speakers who are perceived to be more atypical on the relevant attribute, that is, (perceived) gay speakers for SEXUALITY and women for MASCULINITY. This hypothesis could be tested as a future direction of this work by having listeners only rate male or female talkers.

The weaker indexical role of a retracted /s/ in /stɹ/ clusters may be a result of listeners' experience with the sound change. Many models of sound change propose that a change emerges when listeners stop compensating for extreme coarticulation and instead encode a new speech target (Ohala 1993, Harrignton et al. 2008). Thus for /s/-retraction, the sound change occurs specifically because listeners stop compensating for coarticulation in /stɹ/ clusters. However, the results of this study suggest the reverse pattern – that speakers are compensating more for coarticulation in /stɹ/ clusters than in /s{p, k}ɹ/ clusters, possibly due to their experience with retraction in /stɹ/ clusters. Thus, as speakers are compensating for the retraction in /stɹ/ clusters, /s/ variability is less available to listeners as a sociolinguistics variable in that environment. Conversely, due to their rareness, a retracted /s/ in /spɹ/ or /skɹ/ clusters are not as strongly compensated for and thus can be assigned stronger indexical weight. It may be that this pattern is emerging because most of the listeners are not phonologized retractors and thus may still be expected to perceptually compensate for retraction, and as the sound change progresses, younger generations will compensate less and /stɹ/ variation may take on a stronger socioindexical role.

Furthermore, the prevalence of /s/-retraction in female speakers, whether they are reported to lead the sound change (Wilbanks 2017) or participate at equal levels (Gylfadottir 2014), may also contribute to the weakening of associations of masculinity and sexuality with a retracted /s/ in /stɹ/ clusters, but not in /spɹ/ or /skɹ/ clusters. Future examinations of production data may shed light on whether a gender effect is in fact observed in /s{p, k}ɹ/ clusters, with men exhibited greater numbers of retracted forms, albeit not to the degree of /stɹ/.

5 Conclusion

The results of this study suggest that a retracted /s/ in /spɹ/ and /skɹ/ clusters indexes a more masculine and straight (for men) or lesbian persona, as expected from previous research on a retracted /s/ in prevocalic environments. However, in contrast, a retracted /s/ in /stɹ/ clusters has weaker indexical associations with these attributes, especially in individuals perceived to be less typical for a given attribute. This suggests that /s/-retraction in /stɹ/ clusters is not strongly indexing masculinity, straightness (for men), or lesbianness. This study also found a tremendous degree of individual variability both with respect to how listeners evaluated the attributes and how the talkers were evaluated. Future studies that include more talkers with more robust social information about those talkers may shed more light on the socioindexical role of /s/-retraction.

REFERENCES

BAKER, ADAM, DIANA ARCHANGELI, and JEFF MIELKE. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23. 347–374.

BASS, MICHAEL. 2009. Street or shtreet? investigating (str-) patalisation in Colchester English. Estro: Essex Student Research Online 1. 10–21.

CAMPBELL-KIBLER, KATHRYN. 2011. Intersecting variables and perceived sexual orientation in men. *American Speech* 86. 52–68.

CLARKE, SANDRA. 2004. A note on several unusual fricative pronunciations on the southwest coast of Newfoundland. Regional Language Studies... Newfoundland 18. 15–17.

DARWIN, CHRIS. 2005. Digital mixing script. University of Sussex.

DURIAN, DAVID. 2007. Getting ʃtronger every day?: More on urbanization and the sociogeographic diffusion of (str) in Columbus, OH. *University of Pennsylvania Working Papers in Linguistics* 13. 65–79.

GLAIN, OLIVIER. 2014. Introducing contemporary palatalisation. In T. Lee (ed.), Proceedings of the First Postgraduate and Academic Researchers in Linguistics at York Conference, 16–29.

GYLFADOTTIR, DUNA. 2015. Shtreets of Philadelphia: An Acoustic Study of /stɹ/-retraction in a Naturalistic Speech Corpus. *University of Pennsylvania Working Papers in Linguistics* 21. 2–11.

HARRINGTON, JONATHAN, FELICITAS KLEBER, and ULRICH REUBOLD. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern

British: an acoustic and perceptual study. *Journal of the Acoustical Society of America* 123. 2825–2835

HINRICHS, LARS, ALEXANDER BERGS, AXEL BOHMANN, ERICA BROZOVSKY, BRIAN P. HODGE, KIRSTEN MEEMANN, and PATRICK SCHULTZ. 2015. Sibilants and ethnic diversity: A sociophonetic study of palatalized /s/ in STR clusters among Hispanic, White, and African-American speakers of Texas English. Paper presented at the 16th Texas Linguistic Society conference, University of Texas at Austin.

LAWRENCE, WAYNE P. 2000. /stɹ/ → /ʃtɹ/: Assimilaton at a distance? American Speech 75(1). 82–87.

LEVON, EREZ. 2014. Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society* 43. 539–566.

KRALJIC, TANYA, SUSAN E. BRENNAN, and ARTHUR G. SAMUEL. 2008. Accommodating variation: Dialects, idiolects, and speech processing. *Cognition* 107. 54–81.

OHALA, JOHN. 1993. The phonetics of sound change. *Historical linguistics: Problems and perspectives*, ed. by Charles Jones, 237–278. London: Longman.

PHARAO, NICOLAI, MARIE MAEGAARD, JANUS SPINDLER MØLLER and TORE KRISTIANSEN. 2014. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43. 1–31.

PODESVA, ROBERT J. and JANNEKE VAN HOFWEGEN 2014. How Conservatism and Normative Gender Constrain Variation in Inland California: The Case of /s/. *University of Pennsylvania Working Papers in Linguistics* 20. 129–137.

RUTTER, BEN. 2011. Acoustic analysis of a sound change in progress: The consonant cluster /stɹ/ in English. *Journal of the International Phonetic Association* 41. 27–40.

SHAPIRO, MICHAEL. 1995. A Case of Distant Assimilation: /stɹ/ → /ʃtɹ/. *American Speech* 70. 101–107.

STUART-SMITH, JANE. 2007. Empirical evidence for gendered speech production: /s/ in Glaswegian. *Laboratory Phonology* 9. 65–86.

STUART-SMITH, JANE, MORGAN SONDEREGGER, MICHAEL MCAULIFFE, RACHEL MACDONALD, JEFF MIELKE, ERIK THOMAS and ROBIN DODSWORTH. 2018. Dialectal and social factors affect the phonetic bases of English /s/-retraction. Paper presented at the 16th Conference on Laboratory Phonology, Lisbon.

WILBANKS, ERIC. 2017. Social and Structural Constraints on a Phonetically Motivated Change in Progress: (str) Retraction in Raleigh, NC. *University of Pennsylvania Working Papers in Linguistics* 23. 301–310.

ZIMMAN, LAL. 2013. Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language and Sexuality* 2. 1–39.

Syllabic Size Restrictions on Verb Reduplication in Brazilian Portuguese

JAMILLÄH S. RODRIGUEZ
*University at Albany, State University of New York**

1 Introduction

Brazilian Portuguese incorporates a process of nominalization in which a verb root undergoes total reduplication to form a noun. For example, *pega* 'catch' reduplicates to create *pega-pega* 'a game of tag'. Reduplicated verbs differ from the larger set of verbs with respect to disyllable preference and initial consonants. In this paper, I analyze this difference with the Null Parse theory (McCarthy & Wolf 2009) and compare the Brazilian Portuguese data to an analogous pattern in Cuban Spanish (Lederer 2003).

Using the SUBTLEX corpus (Tang 2012) and past literature (Araújo 2002; Gonçalves 2004; Sempere 2006), a total of 49 reduplicants were compiled. The properties of the reduplicated verbs in this list were compared to the properties of Brazilian Portuguese verbs overall in the SUBTLEX corpus, and there was a difference in size and shape between the two sets of data (henceforth referred to as the reduplicant corpus and SUBTLEX corpus). In the SUBTLEX corpus, disyllabic verbs are common, but they are not the most frequently occuring. However, in the reduplicant corpus, disyllabic verbs are the most frequent. I propose that there is a size restriction of two syllables in verbs that are reduplicated to account for this difference between the characteristics of verbs overall and verbs that are reduplicated. I use the Maxent Grammar Tool (Hayes & Wilson 2006) to examine the predicted outputs using weighted constraints to restrict the size of verbs with different syllable counts. The results show that there is indeed a size restriction on verbs within the reduplication process, and the lack of reduplicated verbs longer than two syllables can be accounted for by a null parse output.

A novel observation of my study on the reduplication pattern is that while monosyllables and disyllables are exclusively consonant-initial, trisyllables are exclusively vowel-initial. Furthermore, a comparison of the two corpora shows that monosyllables are not un-

*My sincere thanks to Michael Becker, Lee Bickmore, Lauren Clemens, and Maria Luisa de Andrade Freitas for their helpful comments, edits, and support. I would also like to thank the audience of BLS 44.

derattested and are not predicted to be worse than disyllables. The forms that are underrepresented are vowel-initial disyllables and both vowel-initial and consonant-initial trisyllables. Contrary to past literature on Portuguese reduplication (Gonçalves 2004) there were not only disyllabic reduplicants, but consonant-initial monosyllables and vowel-initial trisyllables observed in the data as well.

It is common cross-linguistically for phonological phenomena to be restricted to disyllables or at least to preferentially target them. For example, expletive infixation in English is restricted in that the infix appears to the left of a syllabically binary foot (McCarthy 1982).

(1a) abso-bloody-lútely
fan-fuckin-tástic
to-bloody-géther
im-fuckin-pórtant

(McCarthy 1982)

On a related note, there is a disyllabic word minimum requirement in many languages, including several Australian languages like Uradhi and Yidiɲ (Downing 2006).

There are also restrictions that are only applicable to reduplicative processes, such as a reduplication pattern in Tonkawa. Generally in the language, word-initial syllables are usually heavy. In the case of reduplication, word-initial syllables are restricted to light syllables (Gouskova 2007). When a word with a heavy initial syllable is reduplicated, a segment is deleted and the copied reduplicant becomes a light syllable, as in 2.

(1b) naa.toʔs → na-na.toʔs *'I step on it (repeatedly)'*
sal.koʔs → sa-sal.ke.noʔs *'I pull (repeatedly)'*

(Gouskova 2007)

In addition to the factor of word size, the identity of the word-initial segment is a primary determining factor in the reduplication pattern of Brazilian Portuguese. While monosyllables and disyllables are exclusively consonant-initial, trisyllables are exclusively vowel-initial. This is comparable to a reduplication pattern in Timugon Murut. If the stem begins with a consonant the first syllable will be reduplicated, but if the stem begins with a vowel, the first syllable will be skipped and the second syllable will be copied (McCarthy 2008), as in 3.

(1c) li.mo *'five'* → li.li.mo *'about five'*
a.ba.lan *'bathes'* → a.ba.ba.lan *'often bathes'*

(McCarthy 2008)

To summarize, as we will see in Sections 2 and 3, Brazilian Portuguese reduplication is sensitive to commonly typologically observed properties: dyllabicity and whether the word is consonant or vowel-initial.

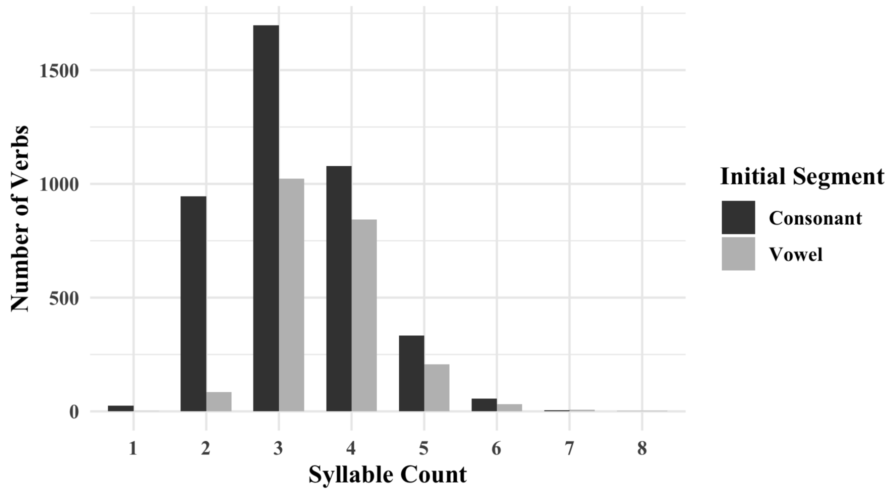
2 Corpus Data

To examine the properties of the Brazilian Portuguese reduplication process, a total of 49 reduplicated verbs were collected from both the past literature and the SUBTLEX corpus. Of the 49 forms, 4 (8%) were monosyllabic, 39 (80%) were disyllabic, and 6 (12%) were trisyllabic. All of the bases were in the third-person singular indicative form. All forms were shared with a native speaker and confirmed to be verbal with a nominal outcome.

In this paper, I use the SUBTLEX corpus to investigate the distribution of syllables and initial segments in Brazilian Portuguese verbs in the overall language. The SUBTLEX corpus consists of approximately 60 million words collected from the subtitles of movies and television shows. Using R (R Development Core Team 2008), I separated the corpus by part of speech, isolated the verbs, and organized them based on syllable count and initial segment. To find the third person verbs, I used the infinitive form, but manually corrected a set of verbs in which the number of syllables in the infinitive differs from the third-person singular indicative form. For example, disyllabic *fazer* 'to have' is conjugated as monosyllabic *faz*.

Figure 1 displays the type frequencies of the third-person singular indicative verbs in the SUBTLEX corpus, separated into consonant-initial and vowel-initial groups. There was a total of 6,340 unique verbs in the corpus.

Figure 1: Type frequencies of consonant-initial and vowel-initial verbs by syllable count in the SUBTLEX corpus



Consonant-initial verbs are more frequent than vowel-initial verbs across all syllable counts. As the syllable count in a verb increases, the difference between consonant-initial and vowel-initial frequency decreases.

To create a realistic model of the reduplication pattern for the Maxent learner, the observed outputs should reflect the proportions that are presented in Table 1.

Table 1: Proportions of predicted (SUBTLEX corpus) and observed (reduplicants corpus) outputs

σ	**Initial Segment**	**Predicted (%)**	**Observed (%)**
1σ	C	0.6	12.0
	V	0	0
2σ	C	25.0	80.0
	V	2.6	0
3σ	C	26.9	0
	V	44.9	8.0

If we look at the ratio of verbs by syllable in the SUBTLEX corpus, we can give the

Maxent learner a predicted output based on observation. In the set of reduplicants that was collected, the ratio of monosyllables to disyllables to trisyllables is 4:39:6. However, based on the corpus of verbs in the overall language, for every 39 disyllabic verbs with an onset there should be 1 consonant-initial monosyllabic verb, 4 vowel-initial disyllabic verbs, 42 consonant-initial trisyllabic and 70 vowel-initial trisyllabic verbs. To account for the difference in the two groups, I added 4 null outputs to the observed ouput of vowel-initial disyllables in the Maxent learner, 42 observed null outputs for the consonant-initial trisyllables, and 64 for the vowel-initial trisyllables.

Although the SUBTLEX corpus contains verbs with more than three syllables, this analysis will focus primarily on monosyllablic, disyllabic, and trisyllabic verbs because there were no observed occurrences of verbs with more than three syllables in the reduplicant corpus. It can be assumed that all stems with more than three syllables do not have a phonetically realized output.

3 Maxent Predictions

In this section, I use the Maxent Grammar Tool to ascertain whether we can accurately portray the reduplication pattern in the language by accounting for the presence of both null and pronounced outputs for disyllables and trisyllables. Although the language consists of verbs with more than three syllables, I will focus only on a maximum of three syllables in the word because that is the limit of the observed reduplication pattern.

In Maxent theory, each constraint is assigned a weight. The harmony ($\mathscr{H}$) is the combined total of the violation marks multiplied by the constraint weight for each candidate, and the probability (p) is the harmony divided by the sum of the harmony of all the candidates.

The null parse candidate, $\odot$, is an output form that holds no phonological or morphological properties. The only constraint that the null parse violates is MPARSE, and as a result, when the null output is present alongside other potential outputs that violate higher-ranked constraints, the null output will always win (McCarhy & Wolf 2009).

I use the following constraints to restrict the size of bases:

(1) PARSE-σ: All syllables must be parsed into feet. (Prince & Smolensky 2004)

(2) FT-BIN: All feet must be syllabically binary. (McCarthy & Prince 1990)

The FT-BIN constraint assigns a violation to candidates with monosyllablic bases due to the absence of a necessary syllable in the binary foot. Because monosyllables and trisyllables have distributional differences in the corpus, the Size-Restrictor constraint, introduced by Junko and Mester 1996, was separated into two constraints: FT-BIN and PARSE-σ rather than one constraint against any non-disyllabic base to account for this output difference.

3.1 Monosyllabic Verbs

Distributional differences between verbs with and without an initial onset were evident within the corpora. As a result, the Maxent learner was given inputs for vowel-initial and consonant-initial verbs separately.

Tableau 2 shows the results of the Maxent learner for monosyllabic reduplication with a verb that has an initial consonant. FT-BIN is only violated by the fully faithful form, and MPARSE is only violated by the null parse candidate.

Table 2: Monosyllabic with word-initial onset, *Dói* 'hurt'

/ˈdɔj/	FT-BIN w=1.2	MPARSE w=11.1	$\mathscr{H}$	p
(ˈdɔj.)(ˈdɔj)	**		-2.4	>**.99**
⊙		*	-11.1	<.01

The fully faithful reduplicated form will most likely be the output form. The null parse candidate has a nearly impossible chance of being the winning candidate.

In the tableau with the vowel-initial input below, the ONSET constraint is added to account for the difference in acceptability of verbs with an initial consonant and verbs without. MPARSE has a high weight, so the null parse candidate should not win, but the combined weights of the other constraints give the phonetically realized candidate a lower harmony.

Table 3: Monosyllabic without word-initial onset, *Há* 'have'

/ˈa/	FT-BIN w=1.2	MPARSE w=11.1	ONSET w=11.9	$\mathscr{H}$	p
(ˈa.)(ˈa)	**		**	-26.2	<.01
⊙		*		-11.1	>**.99**

Tableaux 2 and 3 have inverse results. With consonant-initial verbs, the reduplicated form is almost always the fully faithful, totally reduplicated candidate. However, with vowel-initial verbs the null parse candidate will be the output.

The reduplication of monosyllabic verbs presents a potential problem with the utilization of Null Parse theory, due to the fact that monosyllables are overrepresented in the data and the null parse can only account for underrepresented forms. The proportions of verbs by syllable count and initial segment were re-examined in the Mac-Morpho corpus (Aluísio et al. 2003; Fonseca & Rosa 2013; Fonseca et al. 2015), a corpus of Brazilian Portuguese texts, and were found to be almost identical to the proportions of the SUBTLEX corpus, so the overrepresentation is a consistent issue across corpora and consequently the language.

For disyllabic and trisyllabic data, the observed occurrence for the null parse output was calculated by subtracting the number of pronounced reduplications of that syllable and initial segment observed in the reduplicant corpus from the expected number of verbs according to the SUBTLEX corpus. However, the number of observed monosyllabic reduplications exceeded the number expected based on verbs in the language overall. There were a total of four monosyllablic reduplicants collected, but only one expected. This problem can not be solved using the null parse candidate or the MPARSE constraint.

3.2 Disyllabic Verbs

In the list of reduplicated verbs, disyllabic bases were most common. However, vowel-initial disyllabic reduplicants were absent and only verbs with word-initial consonants were observed. Similar to monosyllables, the MaxEnt Grammar Tool made the following predictions given a disyllabic input.

The fully faithful candidate does not violate any of the constraints and therefore has a harmony of 0. The null parse candidate only violates MPARSE, but because this constraint has a higher weight than 0, the null parse will lose.

Table 4: Disyllabic with word-initial onset, *pega* 'catch'

/ˈpɛ.gə/	FT-BIN w=1.2	MPARSE w=11.1	$\mathscr{H}$	p
(ˈpɛ.gə.)(ˈpɛ.gə)			0	**1**
⊙		*	-11.1	0

If there is a reduplicated verb and that verb is disyllabic and has a word-initial consonant, we expect that it will always be fully faithful and pronounced.

Unlike the consonant-initial verb in Tableau 4, the fully faithful form in Tableau 5 violates the ONSET constraint twice, giving it a low harmony.

Table 5: Disyllabic without word-initial onset, *achar* 'find'

/ˈa.ʃə/	FT-BIN w=1.2	MPARSE w=11.1	ONSET w=11.9	$\mathscr{H}$	p
(ˈa.ʃə.)(ˈa.ʃə)			**	-23.8	$<$.01
⊙		*		-11.1	$>$**.99**

According to the Maxent learner, the predicted probability of a vowel-initial disyllabic reduplicant having a fully faithful output is less than .01%. This differs from Tableau (5) because of the ONSET constraint. Although the MPARSE constraint has a high weight, the addition of the ONSET constraint gives the pronounced candidate a lower harmony than the phonetically unrealized candidate. As a result, the null parse candidate has an almost absolute chance of winning.

3.3 Trisyllabic Verbs

Trisyllabic verb reduplication differs greatly from monosyllabic and disyllabic verb reduplication in Portuguese because of the observed reduplications, both monosyllables and disyllables only had observed occurences of consonant-initial reduplicants. Any vowel-initial reduplicated verbs were assumed to be null pronounced. Contrastingly, only vowel-initial verbs were observed to be reduplicated in trisyllabic verbs.

Furthermore, it was observed that although trisyllabic verbs are most frequent in the SUBTLEX corpus and monosyllables are quite rare, trisyllables are more marked than monosyllables when reduplicated. I add an additional constraint here, LAPSE-MEDIAL, based on constraints in Kager 2001.

(3) LAPSE-MEDIAL: No word-medial unstressed syllables.

Table 6: Trisyllabic with word-initial onset, *carimba* 'stamp'

/ka.(ˈɾĩ.bə)/	PARSE-σ w=0.0	LAPSE-MEDIAL w=23.1	MPARSE w=11.1	$\mathscr{H}$	p
ka.(ˈɾĩ.bə.)ka.(ˈɾĩ.bə)	**	*		-23.1	$<.01$
⊙			*	-11.1	$>$**.99**

As observed in Tableau 6, the null parse candidate will be the output form. The second candidate, [ka.hi.ka.hi.bə], appears to be an ideal candidate. It follows the size restriction requirement for the base and begins with a consonant, which was observed in all of the disyllabic reduplications; but the requirement to remain faithful to the input means it has a small chance of winning.

For monosyllables and disyllables, the ONSET constraint drastically changed the results of the output. With trisyllables, that is not the case.

Table 7: Trisyllabic without word-initial onset, *empurra* 'shove'

/ĩj.(ˈpu.hə)/	PARSE-σ w=0.0	MPARSE w=11.1	ONSET w=11.9	MAX w=1.8	$\mathscr{H}$	p
ĩj.(ˈpu.hĩj.)(ˈpu.hə)	*		*	*	-13.7	.07
⊙		*			-11.1	**.93**

Although the ONSET constraint is introduced to vowel-initial inputs, the output is still most likely to be the Null Parse candidate, with a probability of approximately 93%. However, there is now a 7% chance that the reduplicated form would be the candidate that deletes a segment of the base to avoid a vowel cluster.

Forms in which the base ends in a vowel and the reduplicant begins with a vowel are avoided in Brazilian Portuguese. We can look at verbs with a different initial vowel and final vowel, like [ĩj.ˈpu.hə]. If the base vowel is deleted, we should expect [ĩj.ˈpu.hĩj.ˈpu.hə]. If the reduplicant vowel is deleted, we expect [ĩj.ˈpu.həj.ˈpu.hə]. If the vowels are merged, we should expect something similar to [ĩj.ˈpu.hẽj.ˈpu.hə]. The front, central vowel, [ẽ], is the mid-point between [i] and [ə], but since this is not the observed vowel in the reduplicated form, we can assume that the high, front [i] and the central [ə] are not merging. The data presents the reduplicated form as [ĩj.ˈpu.hĩj.ˈpu.hə].

3.4 Verbs with more than three syllables

Although verbs with more than three syllables are not observed in the reduplication pattern of Portuguese, the probabilities of the candidates for an input can still be predicted. If the input contained a word-initial onset, it is most likely that the output would be null. If the input did not contain a word-initial onset, it is most likely that the output would be null as well, but there would be a small chance that the output could be a phonetically represented form with a deleted segment from the base. This is most similar to trisyllables due to the separation of the size restriction constraint into FT-BIN and PARSE-σ.

4 A Comparison to Reduplication in Cuban Spanish

The Portuguese phenomenon discussed here is similar to a pattern found in Cuban Spanish, which also nominalizes verbs. The resulting noun has a meaning that indicates a repetitive action of the verb (Lederer 2003). For example, *come* ‘eat’ becomes *come-come* ‘a lot of eating’.

Compared to the Cuban Spanish pattern, the reduplication process in Brazilian Portuguese is less restricted in regards to syllable constraints on the length of the base.

Tableau 4 shows the ranking of constraints in Spanish.

Table 8: Verb reduplication in Cuban Spanish, *escucha* 'listen'

/es.ˈku.tʃa/	PARSE-σ	FT-BIN	MPARSE
es.(ˈku.tʃes.)(ˈku.tʃa)	*!		
(es.ˈku.tʃes.)(ˈku.tʃa)		*!	
☞ ⊙			*

The PARSE-σ constraint will reject trisyllabic stems with the condition that all syllables must belong to a foot, and the FT-BIN constraint mandates that there can not be more than two syllables in that foot. PARSE-σ and FT-BIN are ranked higher than the MPARSE constraint, so although the null output, ⊙, violates the MPARSE constraint, it is the winner. The ranking of the constraints in Spanish is PARSE-σ, FT-BIN >> MPARSE.

Table 9: Verb reduplication in Brazilian Portuguese, *agarra* 'grab'

/ə.ˈga.hə/	MPARSE	FT-BIN	PARSE-σ
☞ ə.(ˈga.hə.)(ˈga.hə)			*
(ə.ˈga.hə.)(ˈga.hə)		*!	
⊙	*!		

The ranking of the two constraints in Brazilian Portuguese is the inverse of Spanish: MPARSE >> FT-BIN >> PARSE-σ. Although the fully faithful output violates the size restriction, it will win over the unpronounced candidate.

Although the variance in ranking can account for the difference of outputs between Cuban Spanish and Brazilian Portuguese, it cannot explain why the output in Portuguese is sometimes null for monosyllables and trisyllables and sometimes pronounced. Therefore, I utilize a Maxent framework to give the constraints weights and portray the probabilities of each possible output.

5 Conclusion

In conclusion, there is a size restriction on reduplicated words in Brazilian Portuguese that can be represented by two constraints: FT-BIN and PARSE-σ, with PARSE-σ having a lower weight than FT-BIN. Consonant-initial verbs are more acceptable as bases, except in the case of trisyllables when vowel-initial bases are more acceptable.

Our understanding of the process of verb reduplication in both Brazilian Portuguese and Cuban Spanish would benefit from extension beyond corpus-based studies to an investigation into well-formedness judgments by native speakers. Of particular interest is the fact that this pattern is not as productive as other forms of nominalization and is extremely restricted in a predictable way. Since vowel-initial monosyllables and disyllables, and consonant-initial trisyllables were not observed in the corpora, an experiment utilizing words from those categories could provide more insight to reduplication phonology.

REFERENCES

ALUÍSIO, SANDRA; PELIZZONI, JORGE; MARCHI, ANA RAQUEL; DE OLIVEIRA, LUCÉLIA; MANENTI, REGIANA and MARQUIAFÁVEL, VANESSA. 2003. An Account of the Challenge of Tagging a Reference Corpus for Brazilian Portuguese. In *International Workshop on Computational Processing of the Portuguese Language,* 110-117.

ARAÚJO, GABRIEL. 2002. Truncamento e reduplicação no Português Brasileiro. *Revista de Estudos da Linguagem*, *10*(1).

BARBOSA, PLÍNIO A. and ELEONORA C. ALBANO. 2004. Brazilian Portuguese. *Journal of the International Phonetic Association*, *34*(2), 227-232.

DOWNING, LAURA J. 2006. *Canonical Forms in Prosodic Morphology*. Oxford University Press.

FONESCA, ERIK ROCHA and JOÃO LUÍS G. ROSA. 2013. *Mac-morpho revisited: Towards robust part-of-speech tagging.*

FONSECA, ERICK ROCHA; JOÃO LUÍS G ROSA; and SANDRA MARIA ALUÍSIO. 2015. Evaluating word embeddings and a revised corpus for part-of-speech tagging in Portuguese. *Journal of the Brazilian Computer Society*.

GOUSKOVA, MARIA. 2007. The reduplicative template in Tonkawa. *Phonology*, *24*(3), 367-396.

HAYES, BRUCE and COLIN WILSON. 2006. *Maxent Grammar Tool* (Computer Software). http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/: UCLA.

ITO, JUNKO; YOSHIHISA KITAGAWA; and ARMIN MESTER. 1996. Prosodic faithfulness and correspondence. *Journal of East Asian Linguistics*, *5*, 217-194.

KAGER, RENÉ. 2001. Rhythmic Directionality by Positional Licensing. In *Fifth HIL Phonology Conference (HILP 5).* Handout.

MCCARTHY, JOHN J. 1982. Prosodic Structure and Expletive Infixation. *Language*, *58*(3), 574-590.

MCCARTHY, JOHN J. 2008. *Doing Optimality Theory: Applying Theory to Data*. Wiley-Blackwell.

MCCARTHY, JOHN J. and ALAN PRINCE. 1990. Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory*, *8*(2), 209-283.

MCCARTHY, JOHN J. and MATTHEW WOLF. 2009. Less than Zero: Correspondence and the null output. In (p. 17-66). Equinox Publishing.

PRINCE, ALAN S. 1980. A Metrical Theory for Estonian Quantity. *Linguistic Inquiry*, *11*(3), 511-562.

PRINCE, ALAN and PAUL SMOLENSKY. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Wiley-Blackwell.

R DEVELOPMENT CORE TEAM. 2008. R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org`

SEMPERE, ANTONIO GARU. 2006. *Conflicting quantity patterns in Ibero-Romance prosody* (Unpublished doctoral dissertation). University of Texas at Austin.

TANG, KEVIN. 2012. A 61 Million Word Corpus of Brazilian Portuguese Film Subtitles as a Resource for Linguistic Research. *UCL Work Pap Linguist*, *24*, 208-214.

Viewpoint aspect and object case in Kʷak̓ʷala and Finnish

KATIE SARDINHA
Independent Scholar[*]

1 Introduction

Smith's (1997) two component theory of aspect distinguishes two types of aspectual information. SITUATION ASPECT concerns the classification of events according to their inherent temporal properties – for instance as states, activities, accomplishments, semulfactives, or achievements – while VIEWPOINT ASPECT concerns the classification of events according to which interval of their total duration is under discussion in a context. The relationship between these two types of aspectual information is characterized by Smith (1997:61) as follows.

> Aspectual viewpoints function like the lens of a camera, making objects visible to the receiver. Situations are the objects on which viewpoint lenses are trained. And just as the camera lens is necessary to make the object available for a picture, so viewpoints are necessary to make visible the situation talked about in a sentence.

Smith proposes three canonical viewpoints: IMPERFECTIVE, which focuses on an interval that is internal to an event and does not include its endpoints, as illustrated in 1a; PERFECTIVE, which focuses on the whole event including both endpoints, as illustrated in 1b; and NEUTRAL, which focuses on the initial endpoint plus one event stage, as illustrated in 1c.

(1) Canonical viewpoints: ('I' = initial bound; 'F' = final bound; '/' = visible; '.' = stage)

a.	**Imperfective**	I .../// ... F	(Smith 1997:73)
b.	**Perfective**	I .... F /////////////	(Smith 1997:66)
c.	**Neutral**	I .	(Smith 1997:81)

This paper is concerned with the expression of viewpoint aspectual information in two genetically unrelated languages, Finnish (Uralic) and Kʷak̓ʷala (Wakashan). In discussions about

* A *walas* thank-you to my Kʷak̓ʷala language consultants, Ruby Dawson Cranmer, Mildred Child, Julia Nelson, Violet Bracic, Lily Johnny, and anonymous, for sharing your language with me. I am also grateful to Line Mikkelsen, Erik Maier, and the audience at BLS-44 for questions and discussion which helped me improve the work. Fieldwork for the project was supported by an Oswalt Endangered Language Grant. All errors are my own.

the relationship between viewpoint aspect and object expression, Finnish is the quintessential example language (Travis 2010:133-4). This is because while viewpoint aspect is not grammaticalized in the verbal morphology of Finnish (Smith 1997:5, 81), in certain environments object case functions to signal a contrast between imperfective and perfective viewpoint (Kiparsky 1998, Travis 2010). The purpose of this paper is to provide some initial evidence that in Kʷak̓ʷala, like in Finnish, object case functions in certain environments to approximate a viewpoint contrast which is the mirror opposite, in terms of semantic markedness, of the corresponding imperfective versus perfective contrast found in Finnish. The viewpoint contrast expressed in Kʷak̓ʷala will be referred to as INITIATION VIEWPOINT versus NON-INITIATION VIEWPOINT. This finding builds upon a claim made in Sardinha (2017) that Finnish and Kʷak̓ʷala's object case systems are semantically mirrored. Kʷak̓ʷala will thereby be shown to instantiate an empirically new – yet not unexpected – way for a language to signal viewpoint aspectual information using object case.

The rest of the paper proceeds as follows: Section 2 provides a brief overview of the semantics of object case in Finnish and Kʷak̓ʷala; Section 3 discusses how object case communicates viewpoint aspectual information in Finnish; Section 4 presents evidence for the same, albeit mirrored, pattern in Kʷak̓ʷala; and Section 5 discusses implications of the reported findings and concludes.

2 Object case in Finnish and Kʷak̓ʷala

Finnish and Kʷak̓ʷala each possess two direct object cases, referred to as PARTITIVE and ACCUSATIVE in Finnish, and INSTRUMENTAL and ACCUSATIVE in Kʷak̓ʷala. In this section, I provide a basic overview of these object case systems from the perspective of Sardinha (2017), who argues that they are semantically mirror images of each other.

In Finnish, accusative case relates an internal argument to an event's final subevent, giving rise to an interpretation of boundedness or telicity (Heinämäki 1984, 1994, Vainikka 1989, Kiparsky 1998,[1] Ritter and Rosen 2000, Kratzer 2004, Borer 2005). This semantic value can be clearly observed in sentences with verbs that allow their object to appear in either object case

[1] Kiparsky's (1998) analysis differs substantially from the others referenced here in that the value of "boundedness" is not equated with telicity, but with gradability. Additionally, Kiparsky analyzes both partitive and accusative as meaningful cases, arguing that partitive case licenses unboundedness (rather than being a meaningless default). I do not adopt this analysis on the grounds that partitive objects implicate but do not entail atelicity, as mentioned below.

(henceforth ALTERNATING VERBS).[2] In 2 for instance, accusative case on the object of *ampua* 'shoot' contributes an entailment that the event is telic, which in this example means that the cow was indeed shot (cf. 5 below).[3]

(2)	Metsästäjä	ampui	vahingossa	**lehmän**
	hunter	shot	accident.in	**cow.ACC**
	'The hunter shot **a cow** (ACC) by accident.' (Heinämäki 1984:156)			

An association between accusative case and telicity is also shown by data like 3. The sentence in 3a with the verb *ravistaa* 'shake' is odd because there is no conventional telic end point for the activity of shaking one's legs which would be compatible with the meaning added by accusative case. However, when the event description is modified by a resultative, accusative case on the object becomes possible, as shown in 3b. This is because the resultative adds an explicit end point to the event description, thereby making it compatible with the semantics of accusative.

(3) a.	*	ravistin	**jalkani**	
		I-shook	**legs-ACC-my**	
		Intended: 'I shook **my legs** (ACC).'		
b.		ravistin	**jalkani**	rennoiksi
		I-shook	**legs-ACC-my**	relaxed.to
		'I shook **my legs** (ACC) so that they became relaxed.' (Heinämäki 1994:215)		

Kratzer's (2004:394) analysis of the accusative-assigning head (here, $F_{[acc]}$) is given in 4.

(4) $$[\![F_{[acc]}]\!] = \lambda R_{<e,vt>}.\lambda x_e.\lambda e_v.R(x)(e) \ \& \ \exists f\,[\text{measure}(f) \ \& \ \forall x'\,[x' \leq f(x) \rightarrow \exists e'\,[e' \leq e \ \& \ R(x')(e')]]]]$$

The accusative-assigning head in 4 relates a direct object referent to the temporal extent of an event by turning it into a 'measuring rod' of the event. The nature of this measuring rod is contextually determined and constrained by the semantics of the verb phrase; for instance, the measuring rod

[2] Finnish (and Kʷak̓ʷala) also possess STRICT VERBS whose objects only ever appear in one case (except in circumstances where the meaningful case is semantically licensed, as it is in 9).
[3] Note that the cow's subsequent death may be implied by 2, but is not entailed by it (Heinämäki 1984:156-7).

in 3b is a scale of leg relaxedness. Via the semantics in 4, an accusative object comes to be interpreted as measuring out the event, and this gives rise to an interpretation of the event as telic.

In contrast to accusative, the Finnish object partitive is a meaningless default case. This default value is once again observable in sentences with alternating verbs, where partitive case on the object implicates, though does not entail, atelicity. Thus 5 allows both culminated and non-culminated readings, unlike its counterpart with accusative case in 2.

(5)	Metsästäjä	ampui	vahingossa	**lehmää**
	hunter	shot	accident.in	**cow.PART**

‘The hunter shot **(at) a cow** (**PART**) by accident.’ (Heinämäki 1984:156)

That 5 only implicates atelicity is indicated in Heinämäki (1984:156) as follows: “...the sentence [5] can be used to describe the activity of shooting, no matter what the result is or, the speaker may not know what happened to the cow, and, therefore can not claim anything stronger than [5].”

In summary, Finnish possesses one meaningful object case – accusative – which relates a direct object referent to the final subevent of an event and gives rise to an entailment of telicity. This case is opposed to a meaningless default case, partitive, the use of which implicates atelicity.

Kʷak̓ʷala is the mirror opposite of Finnish in possessing a meaningful instrumental case (=*s*) which relates an internal argument to an event’s initial subevent (Sardinha 2017). Some objects which are consistently realized with instrumental case include semantic instruments, such as the money in 6, and event participants (other than an event’s initiator) which define the initial bound of the event, such as the bighouse which is ‘left’ in 7.[4]

(6)	kəlxʷʔidsuʔnukʷida					babağʷəme	**sis**		**dala**
	kəlxʷ	-xʔid	-suʔnukʷ	=i	=da	babağʷəm	**=s**	**=is**	**dala**
	buy	-BEC	-INDEF.OBJ	=3DIST	=OST	little.boy	**=INST**	**=3REFL.POSS**	**money**

‘The little boy bought something **with his money** (**INST**).’

[4] Abbreviations used in Kʷak̓ʷala glosses are as follows: 1 ‘first person’, 3 ‘third person’, ACC ‘accusative’, AUX ‘auxiliary’, BEC ‘become operator, momentaneous aspect, inchoative’, CONT ‘continuative’, DET ‘determiner’, DIST ‘distal deictic’, EMB ‘embedding -*a*’, FV ‘final vowel’, GRAD.ADV ‘gradual advancement’, INST ‘instrumental’, INDEF.OBJ ‘indefinite object’, INVIS ‘invisible’, MED ‘medial deictic’, NEG ‘negation’, NMZ ‘nominalizer’, OST ‘ostensive determiner’, PART ‘partitive’, POSS ‘possessive’, PREP ‘preposition’, REDUP ‘reduplicant’, REFL.POSS ‘reflexive possessive’, VER ‘verum focus’, VIS ‘visible’. All data not from an otherwise cited source are from my own fieldwork.

(7)

bəwux̌		Mabələx̌		**sa**		**gukʷdzi**
bew	=ux̌	Mabel	=(ə)x̌	**=s**	**=a**	**gukʷdzi**
leave	=3MED	Mabel	=VIS	**=INST**	**=DET**	**bighouse**

‘Mabel left **the bighouse** (**INST**).’

Core event participants in initial subevents (other than the INITIATOR) are referred to as CO-INITIATORS in Sardinha (2017). Several empirical arguments are provided in Sardinha (ibid.) for a semantic link between instrumental case and the property of being a co-initiator. I will limit my discussion here to one of these empirical phenomena, the Direct Manipulation Alternation.

Typically in Kʷak̓ʷala, direct object referents which undergo any sort of change, such as the snow which melts in 8, are expressed in accusative case (as in 8a) and are ungrammatical in instrumental case (as in 8b).

(8) *Context: Ted’s camping. So he builds a fire and melts some snow over it in a pot to make water for him to drink.*

a.

ləm̓is			yax̌ʔid		**x̌a**		**k̓ʷis**
lə	=ʔm	=(w)is	yax̌	-xʔid	**=x̌**	**=a**	**k̓ʷis**
AUX	=VER	=and.so	melt	-BEC	**=ACC**	**=DET**	**snow**
	qəs		naq̓ideʔ				
	q(a)	=is	naq	-xʔid	=a	=iʔ	
	PREP	=3REFL.POSS	drink	-BEC	=EMB	=NMZ	

‘Then he melted **some snow** (**ACC**) to drink.’

b.

* yax̌ʔidux̌			**sa**		**k̓ʷis**	qəs	
yax̌	-xʔid	=ux̌	**=s**	**=a**	**k̓ʷis**	q(a)	=is
melt	-BEC	=3MED	**=INST**	**=DET**	**snow**	PREP	=3REFL.POSS
	naq̓ideʔ						
	naq	-xʔid	=a	=iʔ			
	drink	-BEC	=EMB	=NMZ			

Intended: ‘He melted **some snow** (**INST**) to drink.’

However, this same class of direct object referents can appear in instrumental case when certain semantic conditions are met. In particular, both object cases are grammatical whenever the direct object referent simultaneously undergoes change and serves as the co-initiator of the event. For instance, these semantic conditions are met in 9 by the ice, which undergoes change (by melting) while simultaneously being directly manipulated by the event’s initiator to bring about a change of state (in this instance, a change in itself), thereby serving as the event’s co-initiator.

(9) *Context: Monica held a piece of ice tight between her palms and melted it.*

yax̌ʔidi			Monica	{**sa**, **x̌a**}
yax̌	-xʔid	=i	Monica	{**=s** **=a** , **=x̌** **=a**}
melt	-BEC	=3DIST	Monica	{**=INST** **=DET** , **=ACC** **=DET**}

ƛ̓ux̌ʷ	lax̌is			ʔiʔəy̓əsu	
ƛ̓ux̌ʷ	la	=x̌	=is	ʔi~	ʔəy̓əsu
frozen	PREP	=ACC	=3REFL.POSS	REDUP~	hand/arm

‘Monica melted **the ice** {**INST**, **ACC**} in her hands.’

This phenomenon of case alternation, referred to as the Direct Manipulation Alternation in Sardinha (2017), is possible in those contexts where an argument undergoing change is directly manipulated by the event’s initiator in the course of its undergoing change. The finding that instrumental case can be semantically licensed in this way demonstrates that instrumental is not merely a semantic case for instruments, but is instead associated with a more abstract meaning, namely one tied to event structure and grounded in initial subevents.

Sardinha’s (2017) analysis of the instrumental-case assigning head (here, $F_{[inst]}$) is in 10.

(10) $[\![F_{[inst]}]\!] = \lambda R_{<e,vt>}.\lambda x_e.\lambda e_v.R(x)(e) = 1$ & x is *Co-initiator*[5] of e

The instrumental-assigning head in 10 relates an internal argument to an event’s initial subevent via the event role Co-initiator.

In contrast to instrumental, the accusative case (=*x̌*) in Kʷak̓ʷala is a meaningless default case. One manifestation of this finding is that unlike in Finnish, accusative objects do not give rise to telicity entailments (Greene 2013, Sardinha 2017). This is shown by the possibility of sentences like 11, in which the culmination of an event described using an accusative-marked object is felicitously negated (compare its English translation, which is infelicitous).

[5] $\lambda x_e.\lambda e_v.x$ is *Co-initiator* of e = (x is a dependent cause of e) ∨ (x defines the initial bound of e) ∨ (x is in the possession of an *Initiator* at the initial bound of e). See Chapter 4 of Sardinha (2017) for discussion.

(11)

hiłʔidox̌da				bəgʷanəma	**x̌ən**		**ka**
hił	-xʔid	=ox̌	=da	bəgʷanəm	**=x̌**	**=ən**	**ka**
fix	-BEC	=3MED	=OST	man	**=ACC**	**=1POSS**	**car**

k̓iʔsm̓ox̌			gʷałox̌da			bəgʷanəm
k̓iʔs	=ʔm	=ox̌	ǧʷał	=ox̌	=da	bəgʷanəm
NEG	=VER	=3MED	finish	=3MED	=OST	man

hiłʔi		**x̌ən**		**ka**
hił	-xʔid	**=x̌**	**=ən**	**ka**
fix	-BEC	**=ACC**	**=1POSS**	**car**

‘The man fixed **my car** (ACC) but didn’t finish fixing **my car** (ACC).’ (Greene 2013:44)

More evidence for accusative case being a meaningless default comes from predicates formed using the dummy root, *ʔəx̌-*. Specifically, Sardinha (2017) reports that accusative objects do not restrict the meaning of monotransitive *ʔəx̌-* predicates, while instrumental objects do (Table 1).

CASE FRAME	POSSIBLE INTERPRETATIONS
ʔəx̌- + INST	‘use something’, ‘wear something’
ʔəx̌- + ACC	‘use something’, ‘wear something’, ‘take something’, ‘obtain something’, ‘do to something’

Table 1: Possible interpretations of monotransitive *ʔəx̌-* predicates

In summary, Kʷak̓ʷala possesses one meaningful object case, instrumental, which relates a direct object referent to the initial subevent of an event, and one meaningless default case, accusative.

Figure 1 represents the mirrored nature of Finnish and Kʷak̓ʷala object case systems.

	Initial bound/subevent	***Final bound/subevent***
Finnish	{partitive}	*accusative*
Kʷak̓ʷala	*instrumental*	{accusative}
	(*…* indicates a meaningful case, {...} indicates a meaningless case)	

Figure 1. Semantic mirroring in Finnish and Kʷak̓ʷala object case systems

Figure 1 shows that in both Finnish and Kʷak̓ʷala, a semantic relationship exists between object interpretation and event structure. The two languages differ, however, in terms of which subevent it is – initial or final – which is associated with a meaningful object case.

3 Object case and viewpoint in Finnish

Finnish lacks overt grammatical viewpoint morphemes (Smith 1997:5, 81). Nevertheless, in environments where either object case is grammatical, the semantic contrast encoded through object case gives rise to an imperfective versus perfective viewpoint contrast (Kiparsky 1998, Travis 2010).[6] The pattern is represented in Figure 2, where the semantic value associated with accusative case and partitive case is shown alongside the type of viewpoint information which arises from each case's use.

	Semantic value	***Viewpoint information***
i. *accusative case*	bounded/telic event	perfective viewpoint
ii. *partitive case*	null	imperfective viewpoint (via implicature)

Figure 2. Object case and viewpoint information in Finnish

For instance, with an alternating verb like *luki* 'read', an accusative object is associated with a perfective interpretation (12), while a partitive object is associated by default with an imperfective interpretation (13).

(12) Terttu luki **kirjan**
Terttu read **book.ACC**
'Tertu read (all) **the book** (ACC).' (Heinämäki 1994:212)

(13) Terttu luki **kirjaa**
Terttu read **book.PART**
'Tertu was reading **a book** (PART).' (Heinämäki 1994:212)

[6] For this reason, Finnish object case alternations have occasionally been compared to alternations in Slavic aspectual marking (e.g. Dahl and Karlsson 1976; Kiparsky 1998).

The use of accusative in 12 results in the event being interpreted as telic, and therefore as an event with a final bound; in this way, the use of accusative case consistently gives rise to a perfective viewpoint. On the other hand, the use of partitive in 13 gives rise to an imperfective viewpoint via implicature; as such, this viewpoint information is defeasible (apparently despite the single translation provided for 13). Heinämäki (1994:213) states the following in support of this point.

> ...[13], with a partitive object, is compatible with a situation where Terttu in fact read the whole book, but, for some reason or other, the speaker did not choose to present the situation as bounded. ... But semantically, the sentence [13] is non-committal as to whether the situation itself had some bound or not. In other words, [13] is a non-bounded situation description.

By default, a listener encountering 13 assumes that the speaker has avoided using the accusative case in order to avoid expressing a perfective viewpoint on the event. The listener assumes, therefore, that the speaker intended to communicate an imperfective viewpoint – unless, that is, this assumption is somehow overruled in context. In this way, imperfective viewpoint arises in 13 pragmatically as a result of the semantic opposition between partitive and accusative case.[7]

In summary, in environments where either object case is grammatical, Finnish object case functions to communicate viewpoint aspectual information. The use of accusative case gives rise to a perfective viewpoint via the semantic value of accusative case, while the use of partitive case gives rise to an imperfective viewpoint via implicature. This implicature arises, moreover, due to the null semantics of partitive case together with the enriched meaning that partitive objects receive as a result of the semantic opposition between partitive and accusative.

4 Object case and viewpoint in Kʷak̓ʷala

Kʷak̓ʷala, like Finnish, does not indicate viewpoint aspect grammatically.[8] Nevertheless, since Kʷak̓ʷala's object case system is semantically the mirror image of the one in Finnish, we might

[7] This markedness pattern is reminiscent of what has been reported for various Slavic languages, where the use of perfective verbs is only licit with complete event descriptions, while the use of imperfective verbs is licit with either incomplete or complete event descriptions (Grønn 2003, Alvestad 2014).

[8] Functionally, the nearest thing to a grammatical perfective marker is *–xʔid* (Greene 2013); however, this suffix does not entail telicity and turns out to be neither sufficient nor necessary for communicating (canonical) perfective viewpoint. The nearest thing to a grammatical imperfective is *–n̓akʷəla*; however, this suffix has a more specific meaning than a canonical imperfective (for instance, adding a meaning of graduality to motion events), and while this suffix is sufficient for expressing imperfective viewpoint, it is not necessary for doing so.

expect that in environments where either object case is grammatical, the semantic contrast encoded through object case will give rise to a viewpoint contrast that in terms of markedness, is the mirror image of the imperfective versus perfective contrast found in Finnish. Indeed, this is the claim I will defend here. This claim is summarized in Figure 3, where the semantic value associated with instrumental and accusative case in K^{w}ak̓^{w}ala is shown alongside the type of viewpoint information which arises from each case's use in context. These viewpoints are termed initiation viewpoint and non-initiation viewpoint, respectively.

	Semantic value	***Viewpoint information***
i. *instrumental case*	co-initiated event	initiation viewpoint
ii. *accusative case*	null	non-initiation viewpoint (via implicature)

Figure 3. Object case and viewpoint information in K^{w}ak̓^{w}ala

Initiation viewpoint focuses on an interval within an event's initial (or initiating) subevent, while non-initiation viewpoint focuses on an interval containing an event's final (or non-initiating) subevent.

The empirical evidence for this viewpoint contrast in K^{w}ak̓^{w}ala takes the form of a bias in how sentences are volunteered, and is therefore more subtle than evidence for the corresponding contrast in Finnish. In particular, there is a tendency in K^{w}ak̓^{w}ala for instrumental case to be volunteered in contexts where the speaker is describing an ongoing event, and a tendency for accusative case to be volunteered in contexts where the speaker is describing an event that has been completed or which has resulted in some salient change of state. An example illustrating these tendencies is given in 18: in order to describe an event in which Katie is in the process of putting soup on the stove, the speaker volunteers a sentence with an instrumental object, 18a, while in order to describe an event in which Katie has just put the pot on the stove, a sentence with an accusative object is volunteered, namely 18b (note that the difference in subject-auxiliary ording in these sentences is not semantically significant.)

(18) [*Context: KS is holding an actual pot and acting out a scenario in the speaker's kitchen. When 18a is volunteered, KS is in the middle of slowly placing the pot onto the stove.*]

KS: "If you saw me, doing it?"
Speaker: "Mhm ['Yes']."
KS: "How would you ask — how would you, um, say, 'Katie's putting the pot on the stove'…?"

a.

ləm̓ux̌			Katiyəx̌		hənxƛ̓ənd		
lə	=ʔm	=ux̌	Katie	=(ə)x̌	hən	-xƛ̓	-xʔid
AUX	=VER	=3MED	Katie	=VIS	hollow.container.upright	-on.fire	-BEC

sa		**sup**	lax̌ʷa				ləǧʷilac̓ix̌	
=s	**=a**	**sup**	la	=x̌	=ʷ	=a	ləǧʷilac̓i	=(ə)x̌
=INST	**=DET**	**soup**	PREP	=ACC	=3MED	=DET	stove	=VIS

'Katie's putting **the soup** (**INST**) on the stove.'

KS: "[…] And now let's say I walk away. [*KS has put the pot on the stove and is actually walking away*.] How would you say, 'The soup is on the stove'…?"

b.

ləm̓ux̌			hənxƛ̓əndux̌			Katie
lə	=ʔm	=ux̌	hən	-xƛ̓	-xʔid	=ux̌	Katie
AUX	=VER	=3MED	hollow.container.upright	-on.fire	-BEC	=3MED	Katie

x̌ʷa			**supix̌**		lax̌ʷa			
=x̌	**=ʷ**	**=a**	**sup**	**=(ə)x̌**	la	=x̌	=ʷ	=a
=ACC	**=3MED**	**=DET**	**soup**	**=VIS**	PREP	=ACC	=3MED	=DET

ləǧʷilac̓ix̌	
ləǧʷilac̓i	=(ə)x̌
stove	=VIS

Speaker: "Katie has put the **soup** (**ACC**) on the stove."

Examples 19 and 20 illustrate the same association, this time using consecutive sentences containing the same verb root. In 19, the verb *qəp-* 'pour, spill' takes an instrumental object when the process of pouring is described, as in 19a, but an accusative object when the endpoint of this pouring event is explicitly mentioned, as in 19b. Similarly in 20, the verb *dənx̌-* 'sing' takes an instrumental object when describing the action of singing a song, as in 20a, but an accusative object when referring specifically to the endpoint of this same event, as in 20b.

(19) *Context: Eddie has a bucket with some water, and there's a dog's bowl on the ground.*

a. qəpc̓udi Eddie **sa** **w̓ap**
qəp -c̓u -xʔid =i Eddie **=s** **=a** **w̓ap**
spill -in -BEC =3DIST Eddie **=INST** **=DET** **water**
la x̌a w̓abac̓i
la =x̌ =a w̓abac̓i
PREP =ACC =DET water.dish
'Eddie was pouring/poured **the water** (**INST**) into a water-bowl.'

b. gəlʔəm gʷał qəpa **x̌a** **w̓ap**
gəl =ʔm ǧʷał qəp -a **=x̌** **=a** **w̓ap**
first =VER finish spill -FV **=ACC** **=DET** **water**
laʔe qut̓axʔidida w̓abac̓i
lə =a =i qut̓ -a -xʔid =i =da w̓abac̓i
AUX =EMB =3DIST full -FV -BEC =3DIST =OST water.dish
'Right when he finished pouring **the water** (**ACC**), the bowl got full.'

(20) *Context: Karen entered a charaoke contest, and started singing O Canada — but halfway through she started to feel sick, and had to stop.*

KS: 'Karen sang O Canada, but she didn't finish it.'

dənx̌əlux̌ Karen **sida** **q̓əmdəm** ***O***
dənx̌ -əla =ux̌ Karen **=s** **=i** **=da** **q̓əmdəm** **O**
sing -CONT =3MED Karen **=INST** **=3DIST** **=OST** **song** **O**
Canada, k̓iʔsλux̌ ǧʷał dənx̌ʔidə**x̌**
Canada k̓iʔs =λ =ux̌ ǧʷał dənx̌ -xʔid **=x̌**
Canada NEG =surprise =3MED finish sing -BEC **=ACC**
'Karen was singing/sang **O Canada** (**INST**), but she didn't finish singing **it** (**ACC**).'

In each of these examples, an instrumental object is volunteered when discussing an ongoing event, while an accusative object is volunteered when discussing an event's completion.

This association between object case and viewpoint in Kʷak̓ʷala is, however, only a bias. Speakers do, in fact, volunteer sentences in which the above associations do not hold. Moreover, when sentences such as those in 18-20 are changed by substituting into them whichever object case was not initially volunteered, speakers consistently judge the resulting sentences to be grammatical and insist that case substitution does not change the literal meaning of such sentences. Moreover, while some speakers do comment that case substitution makes a difference of some sort, they consistently struggle to put this difference into words. This is a very different empirical situation compared to what we see in Finnish, where the difference in meaning between sentences with a partitive object versus an accusative object is effable.

This difference in the clarity of empirical evidence derives from the fact that in K^{w}ak̓^{w}ala, instrumental case is semantically redundant in most of those environments where either object case is grammatical.[9] This is true, in particular, whenever the semantic value of instrumental case is redundant with respect to entailments of the verb phrase (as it is in 18-20). The fact that instrumental case is redundant in these environments, together with the fact that accusative case is meaningless, means that the semantic contrast encoded by object case is neutralized in these environments. This neutralization explains why speakers judge sentence pairs like 18a and 18b to be synonymous and insist that they literally mean the same thing. Yet while case choice in these environments makes no semantic difference, existence of the bias illustrated in 18-20 still suggests that case choice makes an informational difference. This difference is at the core of my claim, which is that case choice functions in these environments to signal a viewpoint contrast.

One proposal for how the communication of viewpoint information comes about in these environments is that case choice triggers a relevance implicature. A relevance implicature could arise in 18a as follows: the listener knows that either case would be semantically possible (based on the type of event being described), and registers that the speaker has chosen instrumental case. The listener knows, moreover, that instrumental case associates an internal argument with an event's initial subevent; this leads them to reason that the speaker must have chosen instrumental case intending to highlight the initial subevent of the event as particularly relevant to the discourse. In this way, the speaker invites the listener to see the event from the point of view of its initial subevent, thereby giving rise to initiation viewpoint. A relevance implicature for 18b would proceed along parallel lines: the listener knows that either case would be semantically possible given the type of event being described, and registers the speaker's choice of accusative case. The listener infers that the speaker chose accusative in order to avoid highlighting the initial subevent and concomitantly, to highlight the non-initial (i.e. final) subevent as more relevant to the discourse. Hence, non-initiation viewpoint is born. A relevance implicature analysis along these lines can explain the existence of a bias in how sentences are volunteered illustrated by 18-20.

[9] The Direct Manipulation Alternation illustrated in 9, in which instrumental case is licensed by contextual information, is an example of an environment where instrumental case adds non-redundant information.

Moreover, since relevance implicatures are defeasible, this proposal can also explain why this pattern is merely a bias and not an inviolable constraint.

5 Conclusion

In this paper I've discussed how in Finnish and Kʷak̓ʷala, two languages which do not mark viewpoint aspect grammatically, object case case functions in certain environments to present situations from a certain point of view. Languages like Finnish, in which some aspect of object expression is associated semantically with final subevents and perfectivity, are familiar within the field of linguistics. Kʷak̓ʷala, however, presents us with an empirically new, mirror opposite type of system, in which a meaningful object case is associated with initial subevents and an aspectual viewpoint grounded in initial subevents, here termed initiation viewpoint. More generally, the findings in this paper show that aspectual viewpoints can focus on either initial or final subevents, and that languages may differ in in terms of which subevent is semantically marked. This parameterization is broadly in accord with Ritter and Rosen's (2000) proposal that languages are divided in terms of whether they grammatically privilege the initial or final bound of events. In conclusion, the findings in this paper widen the scope of inquiry into aspectual systems by showcasing a new, though not unexpected, way for a language to signal information about aspectual viewpoint.

REFERENCES

ALVESTAD, SILJE. 2014. Aspect use in the imperative and the prescriptive infinitive in recipes in Bosnian, Croatian, Serbian and Russian: A comparative account. *Jezikoslovlje* 15.2-3. 307-347.

BORER, HAGIT. 2005. *Structuring sense, Volume II: The normal course of events.* New York, NY: Oxford University Press.

GREENE, HANNAH. 2013. *Verb classes in Kʷak̓ʷala*. Vancouver, BC: University of British Columbia MA thesis.

GRØNN, ATLE. 2003. The semantics and pragmatics of the Russian factual imperfective. Oslo, NO: University of Oslo Ph.D. dissertation.

HEINÄMÄKI, ORVOKKI. 1984. Aspect in Finnish. *Aspect Bound: A voyage into the realm of Germanic, Slavonic and Finno-Ugrian aspectology* (*Proceedings from the First Scandinavian Symposium on Aspectology in Vaasa, Aug. 26-28, 1983*), ed. by Casper de Groot & Hannu Tommola, 153-177. Dordrecht, NL: Foris Publications.

HEINÄMÄKI, ORVOKKI. 1994. Aspect as boundedness in Finnish. *Tense, Aspect, and Action: Empirical and Theoretical Contributions to Language Typology* (*Proceedings of seminars on Verbal Semantics at Odense University in 1986 and 1987*), ed. by Carl Bache, Hans Basbøll, & Carl-Erik Lindberg, 207-233. Berlin, DE: Mouton de Gruyter.

KIPARSKY, PAUL. 1998. Partitive case and aspect. *The Projection of Arguments: Lexical and Compositional Factors*, ed. by Miriam Butt and Wilhelm Geuder, 265-308. Stanford, CA: Center for the Study of Language and Information.

KRATZER, ANGELIKA. 2004. Telicity and the meaning of objective case. *The syntax of time*, ed. by Jacqueline Guéron & Jacqueline Lecarme, 398-425. Cambridge, MA: The MIT Press.

RITTER, ELIZABETH, and SARA ROSEN. 2000. Event structure and ergativity. *Events as grammatical objects: The converging perspectives on lexical semantics and syntax*, ed. by Carol Tenny and James Pustejovsky, 187-238. Stanford, CA: CSLI.

SARDINHA, KATHERINE. 2017. *The semantics of Kʷak̓ʷala object case.* Berkeley, CA: University of California Ph.D. dissertation.

SMITH, CARLOTTA. 1997. *The parameter of aspect* (*Studies in Linguistics and Philosophy 43*), *Second edition.* Dordrecht, NL: Kluwer Academic Publishers.

TRAVIS, LISA. 2010. *Inner aspect: The articulation of VP* (*Studies in Natural Language and Linguistic Theory 80*). Dordrecht, NL: Springer.

VAINIKKA, ANNE. 1989. *Deriving syntactic representations in Finnish*. Amherst, MA: University of Massachusetts, Amherst Ph.D. dissertation.

Japanese -*TE* as a marker of aestheticization in the language of menus

KIYOKO TORATANI

York University[*]

1 Introduction

This paper is concerned with characteristics of Japanese -*TE*, variously described as a 'verbal suffix' (Hasegawa 1996:1), a 'gerund' (Martin [1975] 1991:475), a 'linkage marker' (Ohori 1992), or a 'conjugative particle' (Hasegawa 2014:77), among others. As the term 'verbal suffix' suggests, -*TE* is morphologically bound to a verb (V1). Moreover, it is followed by another verb (V2) or a syntactic unit containing the V2, as schematically represented in 1.[1]

(1) … V1-*TE*, … V2.

The semantic function of -*TE* is complex and has received much attention (e.g. Hasegawa 1996). 2 gives typical examples of -*TE*'s use in a sentence.

(2)	a.	Yasai		ga		kit-**te**	aru.		
		vegetable		NOM		cut-TE	be.NPAST		
		'There are vegetables being chopped.'							
	b.	Jon	wa	terebi	o	mi-**te**	ohuro	ni	haitta.
		John	TOP	TV	ACC	watch-TE	bath	LOC	entered.
		'John watched TV and took a bath.'							
	c.	Jon	wa	rakudai.shi-**te**		gakko	o	yameta.	
		John	TOP	fail-TE		school	ACC	quit.	
		'John failed (in an examination) and quit school.'							

(modified from Hasegawa 1996:90, 164, and 165)

[*] I thank questions and comments from the audience at BLS44, especially Koichi Nishida. I also thank Elizabeth Thompson for her editorial assistance. The remaining shortcomings are, of course, my own responsibility.

[1] The following abbreviations are used for Japanese examples: ACC=accusative, DAT=dative, GEN=genitive, H=honorific, NOM=nominative, NPAST=non-past, PRT=particle, QOT=quotative, and TOP=topic.

In 2a, V1 is *kiru* 'cut' and V2 is *aru* '(lit.) exist'. When these two verbs are linked by -*TE*, the sequence expresses a resultative state of V1's argument *yasai* 'vegetable' (i.e. the vegetables are in the resultative state of being chopped) (for more on -*TE aru*, see Tsujimura 2016), whereas in 2b, -*TE* can be roughly translated as 'and' indicating the event centered on V1 (watching) is sequentially followed by the event centered on V2 (taking a bath) (cf. Hasegawa 2014:202), while in 2c, -*TE* is construed as expressing a reason (i.e. failing is the reason for quitting).

Much less discussed is a usage wherein -*TE* occurs as the last element of the sentence: 3 shows its schematic representation, with 4 giving an actual instance.

(3) … V1-*TE*.

(4)

kon'nani	rippa-ni	nat-**te**.
this.much	respectable-DAT	become-TE

'(You) have become such a respectable figure-TE (and)'

(Ohori 1997:476)

Example 4 is a common comment made when a person sees someone who has turned into a fine-looking individual (e.g. a child with whom the speaker is acquainted is now an adult). As the gloss suggests, the translation is incomplete, and a phrase such as 'which makes me so impressed' (Ohori 1997:476) would have to be supplied to be more faithful to the Japanese original. An utterance like 4, which ends in -*TE*, belongs to what Ohori calls 'suspended clauses' (1995:201) or 'suspended clause constructions' (Ohori 1997:471), referring to constructions with 'clauses that bear markers for subordination [such as -*TE*] … and yet are not accompanied by main clauses' (Ohori 1995:201) (cf. 1). They are observed to typically occur 'in the spoken register, especially conversation' (Ohori 1995:213).[2]

Beyond examples like 4, -*TE* as the final element of the sequence is found in a rather unexpected place: that is, in restaurant menus, exemplified below.[3]

[2] This point is consistent with previous work (e.g. Shirakawa 1991; Saegusa 2006), which deal with examples of -*TE* used in conversation.
[3] This dish comes from an Italian restaurant menu (http://rest.la-vita.co.jp/dinner.html) as listed when last accessed in January 2017.

(5)	Yuzu	no	kureepu	shuzetto	**banira muusu**	**o**	**soe-te**
	yuzu	GEN	Crêpe	Suzette	vanilla mousse	ACC	add-TE

‘Crêpe Suzette of *yuzu* [citrus junos], **adding vanilla mousse**-TE’

The sequence in 5 is a name of a dessert dish, which strangely enough ends in -*TE*. This type of name has recently started to show up on restaurant menus, alongside traditional or conventional dish names (see Section 3.1).

To the best of my knowledge, this usage of -*TE*, that is, -*TE* in the item-final position on restaurant menus, has not yet been discussed. Accordingly, this paper is the first attempt to do so. It argues that (a) the item-final -*TE* on menus is syntactically and pragmatically distinct from the utterance-final -*TE* in conversation, and (b) the -*TE* on menus pragmatically functions as a marker of aestheticization.

The organization of this paper is as follows: Section 2 discusses basic differences between utterance-final -*TE* in conversation and the item-final -*TE* on menus; Section 3 looks at the phenomena with -*TE* on menus more closely; Section 4 offers concluding remarks. To gloss the menu item with -*TE*, ‘-TE’ is added at the end to indicate that the phrase does not sound as complete as the English gloss, such as ‘Crêpe Suzette of *yuzu*, adding vanilla mousse’.

2 Utterance-final -*TE* in conversation vs. item-final -*TE* on menus

There are noticeable differences between the utterance-final -*TE* in conversation and the item-final -*TE* on menus. First, the syntactic structure is distinct. As noted earlier, the unit with -*TE* in conversation such as 4 occurs in a ‘suspended clause construction’, implying that it is a sentence. By contrast, since 5 is a dish name, the entire sequence must be a noun phrase. This is represented in 6, using a simplified version of 5.

(6) [[(main part of the dish)]$_{\text{NP}}$ [(elaboration with -*TE*)]]$_{\text{NP}}$

[[*kureepu*	*shuzetto*]	[*banira*	*muusu*	*o*	___*soe-te*]]$_{\text{NP}}$
Crêpe	Suzette	vanilla	mousse	ACC	add-TE

‘Crêpe Suzette, adding vanilla mousse-TE’

The first part is a noun phrase that expresses the name of the main part of the dish, such as *Crêpe Suzette*. This noun phrase is followed by a phonological break, which, in turn, is followed by another phrase ending in -*TE*. As the syntactic category of this phrase is unclear at this point, it is left unspecified. However, it is likely a nominal phrase,[4] modifying the main part of the dish and providing extra information about it. As it does not restrict the meaning, the -*TE*-marked phrase can be analyzed as the semantic equivalent of a non-restrictive relative clause, that is, the adjunct at the NP level (see Van Valin 2005:222).

Second, the mode of delivery is distinct. The utterance-final -*TE* most frequently occurs in 'the spoken register' (Ohori 1995:213), but the item-final -*TE* on menus appears in a 'written-visual medium' (Zwicky & Zwicky 1980:85), which does not involve a face-to-face, continuous, interactional exchange of messages. This difference in the mode of delivery has a consequence in the attachability of sentence-final particles, which are common in conversation but yield oddity in written-visual media.

Sentence-final particles, alternatively called 'pragmatic particles', 'primarily encode the interactional modality' (Iwasaki 2013:302). Consider 7a and 7b, showing a conversation between A and B.

(7)	a.	A:	Jugyoo	yasumu	no?
			class	skip	PRT
			'Are you going to skip the class?'		

[4] It may be counterintuitive to say that a phrase ending in -*TE* is nominal. However, pre-nominally, the phrase with -*TE* patterns like a nominal phrase in that it requires the genitive particle *no*, as shown below.

(i)	te-zukuri	no	hamu		
	hand-making	GEN	ham		
	'hand-made ham'				
(ii)	ryokan-nado	o	riyoo.shite	no	tabi
	inn-etc.	ACC	use-te	GEN	travel
	'a travel of using Japanese inns etc.'				

(Himeno 1983: 36)

(i) shows that a noun phrase *tezukuri* 'hand-making' requires *no* to modify the head noun *hamu* 'ham'. (ii) contains a -*TE*-marked phrase, but it is followed by *no* when it modifies the head noun *tabi* 'travel', just like the noun phrase in (i). Since the genitive case particle attaches to a nominal phrase, it can be argued that the phrase with -*TE* is also nominal. It is also noteworthy that on menus, the modifying element is juxtaposed to the main part of the menu item, although modification in Japanese is normally pre-nominal as in (i) and (ii). The example below is repeated from 13b.

(iii)	Madai	no	poware	asarigai	no	soosu
	seabream	GEN	poêlé	clam	GEN	sauce
	'Seabream butter-roasted, clam sauce'					

In this example, 'clam sauce' follows the main part of the dish 'butter-roasted seabream'. The relationship between the two noun phrases is contextually arrived at: the seabream is either cooked in the sauce or the sauce is added to the seabream as the finishing touch. The point here is that the menu item ending in -*TE* follows the structure of (iii). This serves as another piece of evidence that the phrase with -*TE* is nominal, as it takes the same slot as *asarigai no soosu* 'clam source', a complex noun phrase.

b.	B:	Tsukare-chat-te	ne.
		exhausted-PERF-TE	PRT

‘(I)’m exhausted, you know?’

(Modified from Ohori (1995:202) by adding *you know?* to the gloss for *ne*.)

In 7a, A asks a question, *Jugyoo yasumu no?* ‘Are you going to skip the class?’, and in response, in 7b, instead of directly answering yes or no, B gives a reason for skipping the class, *Tsukare-chat-te ne*. ‘(I)’m exhausted.’ implying that he will skip the class. Here, *-TE* marks the verbal phrase, followed by the sentence particle *ne*, ‘a marker of shared information’ (Iwasaki 2013:303) translated by ‘you know?’ in this context.

By contrast, dish names cannot have *ne* after *-TE*, as in 8.

(8)	#Yuzu no	kureepu	shuzetto	banira muusu o	soe-te	ne
	yuzu GEN	Crêpe	Suzette	vanilla mousse ACC	add-TE	PRT

‘Crêpe Suzette of *yuzu* [citrus junos], adding vanilla mousse-TE, you know?’

It sounds odd to end the menu item with the sentence particle *ne*. 8 reads as if it is part of a casual conversation, where the menu writer is trying to explain something about the menu. As Zwicky and Zwicky (1980:88) note, ‘It is the nature of a menu to be a catalog, a sort of list’, but 9 is outside the appropriate register, an item on the list.

Third, the pragmatic function of *-TE* on menus cannot follow that of conversation, argued as the ‘mitigation of assertion’ (Okamoto 1985) and serving to reduce ‘communicative risks’ by ‘the reduction of the force of speech acts’ (Mihatsch 2013:1). Okamoto (1985:131) states, ‘Verbal and clausal ellipsis [including a sequence ending in *-TE*] is a widely used means for mitigation of assertions in Japanese’, discussing an example like the one below.

(9) A: Odekake desu ka.

‘Are (you) going somewhere?’

B: Ee. Hisashiburini musume no tokoro e itte koyoo to omoimashite.

‘Yes, (I) thought (I) would visit (my) daughter’s place after a long absence *te*.’

(Okamoto 1985:147)

Here, B responds to A's question by ending her utterance with -*TE*. As Okamoto (1985:147) explains, 'It [the use of -*TE*] leaves the addressee more with the feeling that the conversation is to continue, and hence softens the utterance.'

To elaborate the point on 'softening', Saegusa (2006) notes that the utterance-final -*TE* in conversation is often paired with an expression, such as *toka* 'etcetera', and *te-shimau* 'regretfully, shame on me' as in 10, to 'soften' the speaker's claim or statement.

(10)	a.	Sorosoro		nee,	toka	it-te.
		about.time		PRT	QUOT [etc.]	say-TE
		'"It's about time", saying a thing like that.'				
	b.	Yaa,	mainichi		konna,	shokora-shi-te-shimat-te.
		oh	every.day		like.this	chocolate-do-TE-put[regretfully]-TE
		'Oh, I am eating chocolate like this every day, and shame on me.'				

(Saegusa 2006:24)

In 10a, it is grammatical to use a regular quotative particle *to* (as in … *to it-te* 'saying (that)'). Instead a phrase meaning 'etcetera', *toka*, is used to bring about the softening effect. Similarly, in 10b, although it sounds somewhat incomplete, it is possible to end the phrase with *shokora-shi-te* 'doing (eating) chocolate', but instead the speaker adds '*te-shimau*' 'regretfully, shame on me' which softens the utterance.

In contrast, the function of -*TE* on a menu cannot be to soften the message for the purpose of reducing communicative risks, as menus are a 'written-visual medium' (Zwicky & Zwicky 1980:85), involving no immediate communicative risks on a par with an interactional conversation. Furthermore, the additional 'softening' elements such as *toka* 'etcetera' and '*te-shimau*' 'regretfully, shame on me' are inappropriate on menus.

(11)	a.	#Yuzu no	kureepu	shuzetto	banira	muusu	toka	soe-te
		yuzu GEN	Crêpe	Suzette	vanilla	mousse	QUOT[etc.]	add-TE
		'Crêpe Suzette of *yuzu*, adding vanilla mousse, and things like that-TE.'						
	b.	#Yuzu no	kureepu	shuzetto	banira	muusu	o	soe-te-shimat-te
		yuzu GEN	Crêpe	Suzette	vanilla	mousse	ACC	add-TE-put[regretfully]-TE

'Crêpe Suzette of *yuzu*, adding vanilla mousse (and shame on us)-TE.'

Adding *toka* 'things like that' 11a makes the description of the dish imprecise, and this goes against the function of the menu: that is, to inform diners of the specifics of the dish; '*te-shimau*' 'regretfully, shame on me' 11b is read as expressing the chef's regret that s/he added 'vanilla mousse', and this is contradictory to serving vanilla mousse. In other words, the idea of 'softening' cannot simply be extended to account for why *-TE* appears as part of a dish name because of the difference in mode of delivery: conversation may involve immediate communicative risks, thereby needing a softening expression to avoid the risks, but a dish name is in a 'written-visual medium' (Zwicky & Zwicky 1980:85), whose aim is to 'advertise' (Zwicky & Zwicky 1980:86). It is therefore contradictory to include a 'softening' expression when the menu writers aim to convince the diners that they want to have the dish.

The question remains: why is *-TE* used in menus? The next section argues that *-TE* functions pragmatically as a marker of aestheticization, making the dish sound enticing, and high-class. Before discussing the characteristics of *-TE* on menus, it would be helpful to know some basic facts about how conventional Japanese menus entries are presented.

3 Characteristics of item-final-*TE* on menus

3.1 Note on Japanese menus

Traditional Japanese restaurants usually present dish names as short compound nouns, followed by the price, as in 12.

(12) Menu items in traditional Japanese restaurants

焼き鳥	￥250	[*Yakitori*	'grilled chicken on skewers'	250 yen]
月見うどん	￥650	[*Tsukimi-udon*	'Noodle with an egg'	650 yen]
天ぷら定食	￥1,500	[*Tenpura-teishoku*	'Tempura dinner set'	1,500 yen]

In Western restaurant menus, the dish names tend to have a more complex structure as in 13.

(13) Menu items in Western restaurants

a. ムール貝の黒胡椒蒸し ¥900

Muurugai	no	kurokoshoo-mushi
mussles	GEN	black.pepper-steamed

'Mussels steamed with black pepper'

b. 真鯛のポワレ あさり貝のソース ¥2,000

Madai	no	poware	asarigai	no	soosu
seabream	GEN	poêlé	clam	GEN	sauce

'Seabream butter-roasted, clam sauce'

In both instances, the menu items are presented as complex noun phrases with the genitive case particle *no*. 13b is more complex than 13a; it consists of the main part of the dish name (the butter-roasted seabream), followed by a typeset space and additional information about it (clam sauce), implying that the sauce is added to or cooked with the seabream.

It is important to note here that the sequence, *asarigai no soosu* 'clam sauce' in 13b is part of the name of the dish, as proved by the parallel structure of the menu item in 13a. That is, 'clam sauce' is not a secondary description often found in North American menus, illustrated below using a constructed example.

(14) Typical menu entries in North American restaurants

Caesar Salad

$12.00

Romaine, Caesar dressing, Parmesan, olive oil, lemon juice, bacon bits, and croutons.

Avocado Salad

$15.00

Avocado, cucumber, tomato, feta cheese, lemon juice, olive oil, and parsley.

As presented in 14, North American restaurant menus typically consist of three parts: the name of the dish, the price, and a secondary description of the dish, often a list of ingredients. Notably, the name of the dish, such as *Caesar Salad*, appears in large font followed by the price, which is, in turn, followed by the secondary description, with the latter two in smaller fonts.

In the Japanese examples in 13b, the entire sequence, from *madai* ‘seabream’ up to *soosu* ‘sauce’, constitutes the name of the dish, as indicated by the same typeface and the same font size, with the price appearing at the end of the sequence. This characteristic applies to the representation of the menu items with -*TE*, implying that the sequence ending in -*TE* is not a secondary description of the type found in North American menus but is part of the dish name. For instance, the restaurant which has 5 on its menu presents the information as follows: 15a,b give the Japanese original and 15a′,b′ give their respective romanization.

(15)

a.	栗とチョコレートのクッキーシュー	600円	[600 yen]
b.	柚子のクレープシュゼット バニラムースを添えて	700円	[700 yen]

a′.

Kuri	to	chokoreeto	no	kukkii-shuu
chestnut	and	chocolate	GEN	cookie-cream.puff

‘cookie cream puff with chestnuts and chocolate’

b′.

Yuzu	no	kureepu	shuzetto	banira muusu	o	soe-te
yuzu	GEN	Crêpe	Suzette	vanilla mousse	ACC	add-TE

‘Crêpe Suzette of *yuzu* [citrus junos], adding vanilla mousse-TE’

15a is the name of a dessert dish phrased in a complex noun phrase with the genitive case particle *no*, followed by the price. The structure of 15b, repeated from 5b, is parallel to 15a, with the name first, followed by the price. The difference is that the name in 15b is more complex, with the -*TE*-marked phrase after the main part of the dish name, followed by the typeset space. Given this parallelism, it is assumed that the entire sequence up to -*TE* is the name of the dish.[5]

3.2 History

It is unknown precisely when the use of -*TE* on menus started or who started it, but this phenomenon has stirred the curiosity of restaurant menu readers on the Internet. Entries on a Japanese question-and-answer website indicate the phenomenon started sometime in early 2000 in

[5] Some restaurants use a menu structure similar to North American ones; the dish name is in larger font, followed by a description in smaller font with the price. It is possible to find -*TE* in such a secondary description but the present focus is on -*TE* as a part of the primary dish name.

French restaurants.[6] This is in line with a celebrated chef's account of naming French cuisines in an interview dated May 19, 2014.[7] Chef Masayuki Okuda explains that menu writers of French restaurants often give long dish names so that the diners can more easily imagine the content of the unfamiliar dishes simply by reading the name of the dishes. (Recall the contrast in the length of traditional Japanese dish names and Western dish names in 12 and 13). According to Chef Okuda, a typical French dish name starts with the name of the main ingredient, followed by a description of how it is cooked and including a name of the sauce. In the example given on the website, the dish name ends in -*TE* as below, but the chef does not give any account of the usage of -*TE*:

(16) shamo no sotee egoma no miso soosu Suzuki-noojoo no yasai.tachi o soe-te
game.cock GEN sauté perilla GEN miso sauce Suzuki-farm GEN veggies ACC add-TE
'Game cock, sautéed, miso sauce with perilla, adding vegies from Suzuki farm-TE'

Apparently, the use -*TE* on menus emerged as part of this nomenclatural elaboration in French restaurants.

3.3 Variety

Today, the use of -*TE* on menus is not limited to the base verb *soeru* 'add' or to French restaurants. While *soeru* 'add' is the most frequently encountered base verb, other verbs also serve as the base verb for -*TE*, as exemplified below.[8]

(17) a. … fittochiine parumijaano o kezut-te
fettuccine parmigiano ACC shred-TE
'fettuccine, shredding parmigiano (on it)-TE'

b. tamago to buta no oisutaa-itame haru-kyabetsu o tsukat-te
egg and pork GEN oyster-fry spring-cabbage ACC use-TE
'egg and pork, fried in oyster sauce, using spring cabbage-TE'

---

6 The URL is: https://detail.chiebukuro.yahoo.co.jp/qa/question_detail/q1161502311.
7 This comes from a blog entry: http://www.nhk.or.jp/ashita-mirai-blog/187919.html.
8 The menu items come from: 17a: http://www.r-lastricato.com/ristorante/index.html, 17b: http://www.sisen.jp/nagoya/lunch/, 17c: https://www.rihga.com/hiroshima, and17d: http://te-rra.tokyo/menu.

c. fondanshokora… haato no katachi o chiribame-te
fondant.au.chocolat heart GEN shape ACC garnish-TE
'Fondant au chocolat …, garnishing (it) with heart-shaped (cookies)-TE'

d. honjitsu no nama-gaki tappuri kyabia o nose-te
today GEN raw-oyster plenty caviar ACC put-TE
'today's fresh oyster, putting plenty of caviar-TE'

A quantificational study is called for to make a generalization about the semantic and morphological characteristics of the base verbs, but the examples in 17 suggest they should be bivalent (*kezuru* 'shred', *tsukau* 'use', *chiribameru* 'garnish (lit. stud)', *noseru* 'put'), requiring an agent (i.e. the chef) and an object toward which the agent acts, such as an ingredient in the main part of the dish (the cabbage to be used in 17b), or an element added to the main part of the dish (the cheese in 17a, the heart-shape cookies in 17c, and caviar in 17d). Furthermore, these verbs appear in not only different types of dishes, from *a la carte* 17a,d to a main dish 17b or dessert 17c, but also different types of cuisines: Italian 17a, Chinese 17b, French 17c, and Japanese (modern) 17d. In other words, the use of -*TE* is neither an isolated incident restricted to *soeru* 'add' nor to the menu items in French restaurants, but is more pervasive, creating a new trend in menu writing.

3.4 Pragmatic function of -*TE* on menus.

The next question is: why did this trend start? If the sole purpose is to better inform diners about the content of the menu, there is no need to use -*TE* because it can be achieved by the nominal form of the verb, used in menus long before -*TE*. For instance, based on the verb *soeru* 'to attach', the idea of 'attachment' can be expressed either by its nominalized form -*zoe* 'attaching/attachment' as in 18 or by -*TE* (repeated from 5) as in 19.

(18) Yuzu no kureepu shuzetto banira muusu-zoe
yuzu GEN Crêpe Suzette vanilla mousse-addition
'Crêpe Suzette of *yuzu* [citrus junos], with an addition of vanilla mousse'

(19) Yuzu no kureepu shuzetto banira muusu o soe-te
yuzu GEN Crêpe Suzette vanilla mousse ACC add-TE
'Crêpe Suzette of *yuzu* [citrus junos], adding vanilla mousse-TE'

Denotationally, there is no difference between the two menu items. Both denote Crêpe Suzette served with vanilla mousse.[9] But connotationally, the two differ in two respects.

The first is the construal of the agent. The menu item in 18 with the nominal form *-zoe* 'addition' provides an objective description of the state of the dessert with mousse. In other words, the presence of the chef is completely backgrounded, presumably because the referent (chef) remains 'inactive' in the sense of Chafe (1987:25): that is, at the moment of reading the menu item, the referent remains in the menu reader's 'long-term memory, neither focally nor peripherally active'. But 19 implies something extra, engendering the sense of the agent who prepares the dish (i.e. the chef). With *-TE*, the status of the referent (chef) must be 'semi-active'. That is, the presence of chef 'is in a person's [i.e. the menu reader's] peripheral consciousness, a concept of which a person has a background awareness, but which is not being directly focused on' (Chafe 1985:25).

The second point is the event associated with the serving of the dish. Use of *-zoe* 'addition' entails the presence of an event of adding. Use of *soe-te* 'add-*TE*' not only entails the presence of an event of adding but implies presence of another event, although we do not know exactly what it is. This comes from the nature of *-TE*, a linker connecting the base verb to another verbal element (cf. 1). As Ohori (1997:476) argues, *-TE* invites an 'inference-intensive' reading,[10] whereby the hearers (or in this case, the menu readers) can come up with a contextually limited number of phrases as candidates for the unmentioned verbal phrase that follows *-TE*. In 19, the phrases are likely to be those conveying the chef's making or serving of the dish in deference to the customers, shown in 20a and 20b.

(20) a. O-dashi-itashimasu.
H-serve-HUMBLE.NON-PAST
'(I) will (humbly) serve.'
b. O-tsukuri-itashimashita.
H-make-HUMBLE.NON-PAST
'(I) (humbly) made.'

---

[9] The nominal form appears much more frequently than *-TE*. Both forms are sometimes used in one menu.

[10] According to Ohori (1995:213), 'Markers for concession and reason are typically inference-intensive …, and with these markers, the hearer tries to infer what is coming after the dependent clause more actively than otherwise'.

This unsaid portion is the source of the aesthetic sense of elegance and charm, which cannot be yielded by the nominalized form of the verb, such as 18.

This special sense is corroborated by a blogger's careful observation of menu items that end in -*TE*:[11]

(21) Setsuzoku-joshi 'te' de owaru, kono dokutoku na hyoogen. Setsuzoku-joshi nano ni setsuzoku-shinai mama owaru kotode kamoshidasareru, nantomo ienai yoin.
'This unique expression ending in the connective particle "TE" -- An indescribable lingering undertones (*yoin*) is engendered by ending the phrase in -*TE* without connecting to another element, despite -*TE* being a connective particle.'

This blogger uses the term *yoin* 'reverberations' which literally refers to the echoes that remain even after a sound is stopped, whose extended sense is commonly used to describe one's lingering memories after experiencing something. *Yoin* is commonly considered one of the notions that lie at the root of the Japanese sense of aesthetics.

4 Concluding remarks

This paper examines the characteristics of -*TE* on menus, arguing that -*TE* functions as a marker of aestheticization, engendering the sense of elegance, charm, and luxuriousness. First, it points out the differences between the utterance-final -*TE* in conversation and the item-final -*TE* on menus, arguing that the characterization of -*TE* in conversation cannot directly apply to -*TE* on menus, because of the difference in the type of register: the former is a spoken register, but the latter is not. Then, the paper turns to the characteristics of -*TE* on menus, discussing how these items differ from those with a nominalized counterpart of the base verb, the denotational equivalent of the form with -*TE*.

Strauss (2005) discusses the aesthetic function of lexical taste descriptors (e.g. *tasty*) in food advertising language. To this, the paper adds that linguistic aestheticization can be achieved syntactically in Japanese, an example of which is the use of -*TE* in the language of menus.

[11] This comes from: http://www.pictosan.com/menu.html, and it seems to be written by a writer named Keiichi Utsumi who maintains the webpage.

A quantificational study will shed more light on the nature of semantic and morphological characteristics of the base verbs, establish how pervasive the use of *-TE* on menus is, and show to what extent the rate is affected by such factors as cuisine type or restaurant class (cf. Jurafsky 2014).

REFERENCES

CHAFE, WALLACE. 1987. Cognitive constraints on information flow. *Coherence and grounding in discourse*, ed. by Russell Tomlin, 21–51. Amsterdam/Philadelphia: John Benjamins.

HASEGAWA, YOKO. 1996. *A study of Japanese clause linkage: the connective -TE in Japanese*. Stanford: Center for the Study of Language and information.

HASEGAWA, YOKO. 2014. *Japanese: A linguistic introduction*. Cambridge: Cambridge University Press.

HIMENO, MASAKO. 1983. Verb -TE form and its function as noun modifiers. *Nihongogakko ronshuu* 10.25–43.

IWASAKI, SHOICHI. 2013. *Japanese*, *revised edition*. Amsterdam/Philadelphia: John Benjamins.

JURAFSKY, DAN. 2014. *The language of food: A linguist reads the menu*. New York: W.W. Norton.

MARTIN, SAMUEL ELMO. [1975] 1991. *A reference grammar of Japanese*. Rutland.: Tuttle.

MIHATSCH, WILTRUD. 2013. Hedges. The encyclopedia of applied linguistics (vol 4), ed. by Carol A. Chapelle, 1–6. West Sussex: Wiley-Blackwell.

OHORI, TOSHIO. 1992. *Diachrony in clause linkage and related issues*. Berkeley, CA: University of California at Berkeley dissertation.

OHORI, TOSHIO. 1995. Remarks on suspended clauses: A contribution to Japanese phraseology. Essays in semantics and pragmatics: in honor of Charles J. Fillmore, eds. by Masayoshi Shibatani and Sandra A. Thompson, 201–218. Amsterdam/Philadelphia: John Benjamins.

OHORI, TOSHIO. 1997. Framing effects in Japanese non-final clauses: Toward an optional grammar-pragmatics interface. *Berkeley Linguistics Society* 23.471–480.

OKAMOTO, SHIGEKO. 1985. Ellipsis in Japanese discourse. Berkeley, CA: University of California at Berkeley dissertation.

SAEGUSA, REIKO. 2006. Hanashi kotoba ni okeru 'te-kei' [Usage of te-form in spoken Japanese]. *Hitotsubashi daigaku ryuugakusei sentaa kiyoo* 9.15–26.

SHIRAKAWA, HIROYUKI. 1991. 'TE-kei' ni yoru iisashi no bun ni tsuite [On the grammar of truncated sentences ending in *-TE*-form. *Hiroshima daigku nihongo kyooiku gakka kiyoo* 1.39–48.

STRAUSS, SUSAN. 2005. The linguistic aestheticization of food: A cross-cultural look at food commercials in Japan, Korea, and the United States. *Journal of Pragmatics* 37.1427–1455.

TSUJIMURA, NATSUKO. 2016. The role of lexical semantics in the reorganization of the resultative construction. *Transitivity and valency alternations: studies on Japanese and beyond*, eds. by Taro Kageyama and Wesley M. Jacobsen, 155–180. Berlin/Boston: De Gruyter Mouton.

VAN VALIN, ROBERT, D. JR.. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

ZWICKY, ANN, and ARNOLD ZWICKY. 1980. America's national dish: The Style of restaurant menus. *American Speech* 55.87–92.

Parametric dependencies result in correct predictions about word accent typology

ALEXANDRE VAXMAN*
University of Connecticut

1 Introduction

It is well known that metrical stress theories tend to be excessively powerful. Accordingly, the ultimate goal of this paper is to propose a descriptively adequate accentual grammar that would generate all, and only, those phonological accent languages that are effectively attested. ('Phonological accent systems' are systems where accent location is phonologically predictable.) I will present here the parametric component of the SCALES-AND-PARAMETERS (S&P) theory (Vaxman 2016), which differs from metrical theory in that it assigns accent ('primary stress') separately from rhythm ('non-primary stress'). This is motivated by a range of asymmetries with respect to their phonological behavior (see Goedemans and van der Hulst 2014, van der Hulst 2010, McGarrity 2003).

2 The parameter system

2.1 The parameters

In S&P, accent is assigned by a parameter system within the ACCENT DOMAIN (the span of syllables where accent may fall in a given language) which is placed on the ACCENT GRID. The parameter system contains 7 binary parameters listed in 1.

(1) The S&P parameters

a. The Domain Size parameter (DS): the accent domain is {Bounded, Unbounded}.

b. The Domain Edge parameter (DE): a bounded accent domain is formed at the {Left, Right} word edge.

* The author is grateful to Harry van der Hulst, Francesc Torres Tamarit, Noam Faust, Nicholas Rolle, Bernard Tranel, and audiences at the BLS44, RFP16 and 24mfm for useful feedback. Special thanks to Anne Carrio, Martine and Robert Maculet, Jean-Marc Henry and EK.

c. The Nonfinality parameter (NF): the peripheral element at the right word edge is not allowed to receive accent. {Yes, No}

d. The Nonfinality Unit parameter (NF UT): the nonfinality unit is a {Syllable, Segment}.

e. The Weight parameter (W): the language has weight distinctions. {Yes, No}

f. The Project Position parameter (PP): project {Leftmost, Rightmost} position in the accent domain onto line 1 of the Accent Grid.

g. The Select parameter (SEL): choose the {Leftmost, Rightmost} gridmark on line 1 by placing a gridmark over it on line 2.

The Domain Size parameter determines the size of the accent domain: bounded (i.e. binary) 2a, or unbounded (i.e. the entire word) 2b.[1] In languages with DS (Bounded), the accent domain is located near a word edge: for example, DE (Left) yields 2a, DE (Right) yields 2c; in those with DS (Unbounded), it is co-extensive with the entire word, modulo an extrametrical (EM) final unit, when NF (Yes) makes the last unit invisible to accent assignment 2d. This unit may be either a syllable or a segment.

(2)	a. [(ˈh l)	ˈsontako	young girl	(Capanahua)
	[(l ˈh)	wiˈrankin	he pushed it	
	b. [(l l ˈh h h l)]	hapˀaˈlaamaubiiwi	mud	(Yana)
	c. (ˈh h)]	ʔaˈsirˌtar	lucky	(Aklan)
	(l ˈh)]	kiˌnapuˈtus	wrap tool-FOC-PAST	
	d. (l ˈh) <σ>]	reˈfeːkit	remake-PERF-3Sg	(Cl. Latin)

[1] By convention, parentheses ‘( )’ stand for the boundaries of the accent domain: thus, (σσ) corresponds to a bounded (binary) domain; square brackets (‘[‘ and ‘]’) represent, respectively, the left and right edge of the word; angled brackets (‘< >’) enclose an extrametrical unit.

Note that S&P does not recognize initial EM, deemed suspect in view of many wrong or incomplete reports, such as the one for Negev Bedouin Arabic in Blanc (1971), and misanalyses, notably in the case of ‘broken window’ systems such as Kashaya.

In words with heavy syllables, these are projected as gridmarks onto line 1 of the Accent Grid (Weight Projection, or ‘WP’); then, the Select parameter chooses one of these gridmarks by placing a gridmark on line 2. This (tallest) gridmark column is ‘read off’ as accentual prominence 3. For example, in bounded systems with the right-edge accent domain, if Select is set to ‘Left’, then the left(most) gridmark in the domain is chosen for accent, resulting in penultimate accent; if Select is set to ‘Right’, then the right(most) gridmark in the domain is chosen, resulting in final accent. Similarly, in unbounded systems, Sel (L/R) assigns accent to the leftmost/rightmost gridmark in the domain, which, in this case, is the entire word.

(3)	DS (**Bounded**)	DS (**Unbounded**)		DS (**Bounded**)	DS (**Unbounded**)		
	*	*	Sel (**L**)	*	*	Sel (**R**)	line 2
	* *	* * *	WP	* *	* * *	WP	line 1
	(ˈh h)]	[(l ˈh l h h l)]		(ˈh h)]	[(l ˈh l h h l)]		

2.2 Relations between parameters

An important aspect of the S&P theory are the dependency relations holding among certain parameters. By dependencies between parameters A and B, I understand (in the traditional sense) a relation whereby, for some value of A, B may **not** be set to at least one of its values. A parameter that may not be set to **any** value is said to be blocked.

Some dependencies in the S&P parameter system are intrinsic (i.e. they entirely follow from the definitions of the relevant parameters). For example, if the entire word forms the accent domain, it makes no sense asking at which word edge this domain is located. We will say that, if DS (Unbounded), then DE may not be set to any value (i.e. it is blocked); therefore, DE is dependent on DS. Similarly, if a language does not have nonfinality, it does not have a NF unit, either. That is, if NF (No), then NF Ut is blocked; therefore, NF Ut is dependent on NF.

In generating a given 'language' (a set of accentual patterns), the parameters apply successively in a fixed **UNIVERSAL** order. (In this respect, S&P is more restrictive than OT where language-specific constraint reranking results in strong overgeneration.) The order in which the parameters apply is often (but not always) implied by parameter dependencies. For example, if NF (No), then NF Ut is blocked. Hence, deciding whether NF Ut may be set requires knowing how NF is set. Therefore, NF is set prior to NF Ut. However, for certain parameters, order of application cannot be inferred from dependencies. For example, Weight and Select are independent, but Weight, clearly, precedes Select because the latter operates on heavy syllables, which requires knowing whether the system is WS, in the first place.

In addition to intrinsic dependencies, the system also contains several EXTRINSIC (i.e. empirically supported) dependency relations. These play a major role in predicting the actual accentual typology by significantly reducing the parameter space in such a way that the theory closely approaches descriptive adequacy.

3 Establishing the Accent Locality Dependency

In this section, based on empirical evidence, I establish one important extrinsic dependency, which I call the ACCENT LOCALITY DEPENDENCY (ALD). The proposal receives empirical support from tests against information in STRESSTYP, the largest-to-date database of stress patterns in the world's languages, implemented and enriched since mid-1990s by Rob Goedemans, Harry van der Hulst and a number of colleagues. The particular version of StressTyp1 which I have used contains records for more than 550 genetically diverse languages. Most importantly for the present purposes, StressTyp cites both primary and theoretical sources, provides examples and offers (preliminary) formal analyses.

3.1 An asymmetry for Select in bounded WS systems

Work with StressTyp has revealed that in **bounded** WS systems where the final **syllable** is **EM**, accent falls on that heavy syllable (if any) which is closest to the right edge of the word. For example, in Cl. Latin, forms like [konˈstruktus] ‘collected together’ are grammatical, while those where accent is more to the left than the last heavy syllable, like *[ˈkonstruktus], are ill-formed.

The StressTyp query for bounded systems with a final EM syllable returns 23 languages (see Table 1), out of which 19 are consistent with the observation above, while 4 are reported to have Select (Left).

TABLE 1. *Bounded WS languages with NF (Yes) and NF Unit (Syllable) in StressTyp.*

NF	NF Unit	Select	Languages
Yes	Yes	**Left**	Roro, Bhojpuri, Central Sierra Miwok, Hopi
Yes	Yes	Right	19 languages

For three among those (Central Sierra Miwok, Bhojpuri, Roro), careful examination reveals that, in fact, these are not counterexamples.

Central Sierra Miwok and Bhojpuri are reported in StressTyp as involving NF, but, in fact, they turn out not to have it. Thus, in C. Sierra Miwok, evidence for the lack of NF comes from forms like [ˈpalat̪:aˌt̪a] ‘woodpecker’, where accent is initial (Freeland 1951). Indeed, nonfinality is logically impossible in a language where accent is assigned at the opposite edge. In Bhojpuri, the final closed syllable containing a short vowel is accented, as in [gaˈlab] ‘to melt’ (Hayes 1981). Since both C. Sierra Miwok and Bhojpuri do not involve NF, they fall out of the scope of the observation above.

According to Hayes (1981), Roro stresses the leftmost heavy syllable in the word. However, this analysis is incorrect because it does not match the (only) primary source (Strong 1913-1914) which notes several regular WI patterns. Moreover, Strong’s ‘Accent’ section completely lacks accentual data and has four lines, only. Thus, given the lack of empirical evidence, Roro does not constitute a counter-example to the observation made above.

As for Hopi, this is the only genuine exception. It is worth noting, though, that all exceptional forms are restricted to a single, highly marked configuration (Jeanne 1978, Kalectaca 1978).

The observed bias towards Select (Left) in bounded systems suggests that a broader generalization may be stated. I submit a strong claim that this observation extends, in fact, to **all** types of WS systems with NF, namely to systems (bounded as well as unbounded) where the NF unit is not the syllable and to unbounded systems with NF.

3.2 The Accent Locality Hypothesis

3.2.1 The statement

The relevant generalization, which I call the ACCENT LOCALITY HYPOTHESIS (ALH), is stated as follows:

(4) If a WS system involves Nonfinality, then in words with heavy syllables, accent falls on that heavy syllable which is closest to the right word edge.

For example, similar to bounded systems, in unbounded ones, NF (Yes) does not combine with Sel (L). This is illustrated by the regular Western Mari form in 5a, as opposed to the ill-formed 5b.

(5) a. oʃˈmaʃtə sand-INESS b. *[(l ˈh h) <l>] (V_{full} heavy, /ə/ light)

Note that the possible accent locations are not limited to the configuration where accent falls on the heavy syllable adjacent to the NF syllable, as in 5a. Rather, in unbounded systems, the accent on the **last heavy** syllable in the accent domain may fall deeper inside the word when this syllable is followed by light ones, as the pattern [(l h h l ˈh **l l**) <σ>] illustrates.

From the ALH, the testable and falsifiable prediction 6 may be derived.

(6) In WS languages with NF, heavy syllables other than the rightmost heavy one in the accent domain do not bear word accent.

In order to falsify 6, it is sufficient to present forms that have accent further to the left of the rightmost heavy syllable in the accent domain in WS systems with nonfinality.

3.2.2 Testing the prediction

Let us now test 6 to verify whether it is true of all types of WS systems with NF.

For bounded WS systems with the NF Unit other than the syllable, the StressTyp query returns languages with Select (Right), but none with Select (Left), as displayed in Table 2.

TABLE 2. *Bounded WS languages with NF units other than the syllable.*

NONFINALITY	*NF UNIT*	*WEIGHT*	*SELECT*	*LANGUAGE*
Yes	Consonant	Yes	Right	Cebuano,Evenki
Yes	Heavy syllable	Yes	Right	Dutch, Norwegian
Yes	Mora	Yes	Right	Hindi
Yes	Echo vowel	Yes	Right	Tobelo

I conclude that the ALH is fully supported for bounded systems with NF.

For unbounded systems with NF, the StressTyp query returns the output in Table 3.

TABLE 3. *Unbounded WS languages with NF (Yes).*

NONFINALITY	*NONFINALITY UNIT*	*SELECT*	*LANGUAGE*
Yes	Syllable (?)	Left	**Zeberio Basque**
Yes	Vowel	Left	**Tahitian**
Yes	Segment	Left	**Gorowa**
Yes	Syllable	Right	Sindhi, Western Mari
Yes	Consonant	Right	Kenuzi-Dongola
Yes	Segment	Right	Cyrenaican Bedouin Arabic
Yes	Foot	Right	Munsee, Unami, Passamaquoddy

Summarizing, there are 7 unbounded WS systems with NF that obey the ALH in Table 3.

Three languages set in bold (Zeberio Basque, Tahitian, Gorowa), reported with NF (Yes) and Select (Left), seemingly run counter the ALH. However, as I will now show, these are spurious counter-examples.

First, Zeberio Basque is, in fact, a lexical accent system with a special accent-attracting role for the root. Since this is not a phonological accent system, it should be discarded.

Second, Tahitian is reported as ‘stress the leftmost heavy’ with the second part of the final vowel being EM (Hayes 1981, after Tryon 1970). But in fact, as established by Bickmore (1995), accent falls on the **rightmost** heavy syllable 7a. Also, Tahitian lacks NF. Indeed, the second part of long vowels and diphthongs was described as EM in the literature. In that case, however, only the left part of the final long vowels in 7b would be visible to accent assignment and, accordingly, the last syllable would behave as light. This makes the wrong prediction that accent falls on a heavy non-final syllable. Given that the final long syllable does get the accent in the representative examples in 7b, I conclude that Tahitian does not have Nonfinality.

(7)	a. ta:ni:ni:ˈto:	be dizzy	a:ˈvo:ta	avocado	ˌpa:ˈrau	oyster
	b. o:ˈpu:	stomach	pa:to:ˈto:	knock		

Since Tahitian has neither NF (Yes), nor Select (Right), it does not contradict the ALH.

Third, while Gorowa has been reported with NF (Yes), ultimate accent in some words, such as [gaˈla] ‘which’ and [aoˈwa] ‘drink’, shows Gorowa does not have NF. Therefore, Gorowa is not counter-evidence to the prediction.

To sum up, based on tests of the ALH against StressTyp data (revised using reliable publications), that the ALH receives strong empirical support.

3.3 The Accent Locality Dependency

The Accent Locality Hypothesis, shown correct above, implies that the ACCENT LOCALITY DEPENDENCY (ALD) in 8 holds true.

(8) [Nonfinality (Yes) & Weight (Yes)] → Select (Right)

In the next section, we will consider typological implications of parameter dependencies.

4 A parametric typology for S&P

The interest of parameter dependencies (such as the ALD) for accentual theory lies in parameter space reduction. Thus, S&P **accurately** reduces the parameter space, so that the system almost reaches descriptive adequacy. The types of WS languages generated by the S&P parameter system are listed in Table 4. As shown, all the predicted languages in this typology are attested.

TABLE 4. *Types of WS languages generated by the S&P parameter system.*

DS	DE	SEL	PP	NF	ATTESTED?
B	L	L	L	No	Yindjibarndi
B	L	L	R	No	Ossetic
B	L	R	L	No	Capanahua
B	L	R	R	No	Archin
B	R	L	L	No	Sundanese
B	R	L	R	No	Aklan
B	R	R	L	No	Epera Pedée
B	R	R	R	No	Yapese
B	R	R	L	Yes	Classical Latin
B	R	R	R	Yes	Cebuano
U	N/A	L	L	No	Au
U	N/A	L	R	No	Kwak’wala
U	N/A	R	L	No	Kuuku Yaʔu
U	N/A	R	R	No	Aguacateco
U	N/A	R	L	Yes	Sindhi
U	N/A	R	R	Yes	Western Mari

In addition to these 16 attested WS language types, four WS types with NF (Segment) are predicted (as the ALD sets no restrictions on the NF unit). This prediction will be tested in a future study.

The S&P parameter system also yields 5 types of weight-**insensitive** languages. It must be said that S&P analyzes all fixed accent systems as bounded. As van der Hulst (1996) notes, fixed (WI) accent systems (at least, in the absence of exceptions) are ambiguous between a bounded and unbounded analyses. In order to avoid parametric ambiguity, problematic for the learner, S&P forces the bounded analysis by ruling out the unbounded one.

In WI languages, syllables do not attract accent: all forms contain light syllables, only. Therefore, in these languages, accent is assigned by the default mechanism which, in S&P, essentially involves the Project Position parameter. Accent falls on one of the two peripheral positions at an edge, as determined by PP, together with DE. In addition, in systems with NF, the extrametrical syllable shifts the accent domain one syllable inside the word, deriving antepenultimate accent with PP (Left), as in Macedonian. All 5 WI systems generated by S&P are attested, as shown in 9.

(9)

Initial	**Second**	**Third**
PP (L), Sel (L)	PP (R), Sel (R)	No left-edge EM
*	*	
*	*	N/A
[(l l)	[(l l)	
Czech	Araucanian	
Final	**Penultimate**	**Antepenultimate**
PP (R), Sel (R)	PP (L), Sel (L)	PP (L), Sel (L), NF (Y)
*	*	*
*	*	*
(l l)]	(l l)]	(l l) <σ>]
Quiché	Polish	Macedonian

Summarizing, the S&P parameter system generates a total of 25 different types of languages, out of which 21 (16 WS, 5 WI) are effectively attested, while 4, with NF Unit (Segment), await future study.

5 Comparing parametric typologies for the S&P and Primary Accent First theories

This section offers a brief comparison between the S&P parameter system presented here and the Primary Accent First (PAF) grammar (van der Hulst 1996, 2010, 2012), from which S&P descends, in terms of their generative capacity and typological fit.

To begin with, let us relate the parameters across the two theories. Since S&P doesn't recognize initial EM, only the 'Right' value of PAF's EM parameter may be equated to NF(Yes) of S&P. Further, PAF's Domain parameter is ternary: it admits Left/Right/not set where 'not set' yields unbounded systems (van der Hulst 2012). Accordingly, PAF's Domain (Left/Right) corresponds, in S&P, to the combination of DS (Bounded) & DE (Left/Right), whereas Domain (not set) corresponds to DS (Unbounded).

Given these correspondences, Vaxman (2016) shows that all 16 WS language types generated by S&P (Table 4) are also generated by the PAF grammar. But, in addition, PAF generates 20 (!) unattested language types (see Table 5).

TABLE 5. *Unattested languages generated by the PAF, but not S&P, grammar*.

DOMAIN	EM	SELECT	DEFAULT	ATTESTED?
Left	Left	Left	Left	Unattested
Left	Left	Left	Right	Unattested
Left	Left	Right	Left	Unattested
Left	Left	Right	Right	Unattested
Right	Left	Left	Left	Unattested
Right	Left	Left	Right	Unattested
Right	Left	Right	Left	Unattested
Right	Left	Right	Right	Unattested
Left	Right	Left	Left	Unattested
Left	Right	Left	Right	Unattested
Left	Right	Right	Left	Unattested
Left	Right	Right	Right	Unattested
Right	Right	Left	Left	Unattested
Right	Right	Left	Right	Unattested
Not set (UNB)	Left	Left	Left	Unattested
Not set (UNB)	Left	Left	Right	Unattested
Not set (UNB)	Left	Right	Left	Unattested
Not set (UNB)	Left	Right	Right	Unattested
Not set (UNB)	Right	Left	Left	Unattested
Not set (UNB)	Right	Left	Right	Unattested

In conclusion, while PAF and S&P both generate the same set of attested languages (i.e. they do not undergenerate), the former massively overgenerates. This result highlights an original feature of the S&P system, that is, incorporation of parameter dependencies, which leads to a significant weakening of its generative capacity. At the same time, the dependencies reduce the parameter space so as to avoid undergeneration.

6 Conclusion

In this article, I introduced and motivated the parametric component of the Scales-and-Parameters theory whose parameter system is designed to account for cross-linguistic variation in accent location, in particular, relating to phonological accent systems.

I described the content and organization of the S&P system, arguing that dependency relations hold among certain parameters (Section 2). The system consists of a set of 7 binary parameters which operate on the Accent Grid, a prosodic representation for word accent assignment. (Rhythm, or 'non-primary stress', is assigned on a separate phonological plane by a different mechanism not considered here.)

The parameters apply in a fixed universal order, which makes the system typologically restrictive (compared to language-specific constraint reranking in OT).

Importantly, certain parameters in the S&P system are dependent on others. Some of those dependencies are intrinsic, as they derive from the content of the parameters themselves, while others (extrinsic dependencies) capture falsifiable empirical hypotheses that have received external support (Section 2.2).

In particular, the Accent Locality Hypothesis (Section 3.2.1) in 4 leads to the Accent Locality Dependency (Section 3.3) in 8 between the parameters Select, Nonfinality and Weight.

(4) If a WS system involves Nonfinality, then in words with heavy syllables, accent falls on that heavy syllable which is closest to the right word edge.

(8) [Nonfinality (Yes) & Weight (Yes)] → Select (Right)

The prediction in 6 (Section 3.2.1), derived from the Accent Locality Hypothesis, has been tested and shown to be borne out, based on information in StressTyp, currently the largest database of stress patterns in the languages of the world. I was led to revise certain records based on many publications, notably the primary sources.

Lastly, the comparison of PAF and S&P by means of a parametric typology (Sections 4-5) has established that, in the case of phonological accent systems, the former strongly overgenerates, while the latter significantly reduces the parameter space in specific ways that bring it close to descriptive adequacy.

REFERENCES

BICKMORE, LEE S. 1995. Refining and formalizing the Tahitian stress placement algorithm. *Oceanic Linguistics* 34:2.410-42.

BLANC, HENRI. 1971. The Arabic dialect of the Negev Bedouins. In *Proceedings of the Israel Academy of Sciences and Humanities* 1:7.112-50.

FREELAND, L. S. 1951. *Language of the Sierra Miwok.* (Indiana University Publications in Anthropology and Linguistics. Memoir 6 of the International Journal of American Linguistics.) Baltimore: Waverley Press.

GOEDEMANS, ROB, and H. G. VAN DER HULST. 2014. The separation of accent and rhythm: Evidence from StressTyp. *Word stress: Theoretical and typological issues*, ed. by H. G. van der Hulst (ed.), 119-45. Cambridge: Cambridge University Press.

HAYES, BRUCE. 1981. *A metrical theory of stress rules*. Cambridge, MA: MIT dissertation.

JEANNE, LAVERNE MASAYESAVA. 1978. *Aspects of Hopi grammar*. Cambridge, MA: MIT dissertation.

KALECTACA, MILO. 1978. *Lessons in Hopi.* Tucson: University of Arizona Press.

MCGARRITY, LAURA W. 2003. *Constraints on patterns of primary and secondary stress*. Bloomington, IN: Indiana University dissertation.

STRONG, W. M. 1913-1914. The Roro and Mekeo languages of British New Guinea. *Zeitschrift für kolonialsprachen* 4.285-311.

TRYON, DARREL T. 1970. *Conversational Tahitian. An introduction to the Tahitian language of French Polynesia.* Berkeley & Los Angeles: University of California Press.

VAN DER HULST, H. G. 1996. Separating primary and secondary accent. *Stress patterns of the world.* ed. by Rob Goedemans, H. G. van der Hulst and Ellis Visch, 1-26, Part 1 (HIL Publications 2). The Hague: Holland Academic Graphics.

VAN DER HULST, H. G. 2010. Representing accent. *Phonological studies* 13.117-28.

VAN DER HULST, H. G. 2012. Deconstructing stress. *Lingua* 122.1494-521.

VAXMAN, ALEXANDRE. 2016. *How to beat without feet: Weight scales and parameter dependencies in the computation of word accent*. Storrs, CT: University of Connecticut dissertation.