On the diachronic nature of Marathi light verbs

Aaditya Kulkarni
The M. S. University of Baroda, India.
aaditya.kulkarni.94@gmail.com

The diachronic nature of light verbs in the Indo-Aryan languages has been a matter of debate. Scholars like Hook (1991, 1993) and Slade (2013) consider light verbs to have emerged as the resultant of diachronic change; whereas Butt and Lahiri (2013) opine that light verbs are historically stable and resistant to change. In the light of these views, the present paper traces the Marathi light verbs across time and tries to show that their pattern is indicative of gradual emergence rather than historical stability.

A light (or a vector) verb is bleached of its lexical meaning which its homophonous full verb counterpart expresses, but adds semantic nuances to the meaning expressed by the main verb when it occurs in a compound verb construction, as illustrated by '*ʈakɳe*' in (1):

1.	*mi*	*pəd̪trə*	*lih-un*	*ʈak-l-e*
	1.SG	letter.N	write-CP	drop-PRF-N

I wrote the letter (got rid of the task/responsibility of writing it).

Light verbs were believed to be absent in old Indo-Aryan, and were considered to have developed during the middle Indo-Aryan period to compensate for the loss of root modifying affixes (Beams, 1879). Taking this forward, Hook (1991, 1993) proposed that the development of light verbs in the Indo-Aryan languages is a case of 'aspectogenesis' (emergence of aspectual contrast). He posited that the difference between a Hindi compound verb and its corresponding simple verb construction (*kərna* vs. *kər dena*) is mainly aspectual, and the usage of a light verb gives a perfective reading as opposed to its simple verb counterpart. In stark contrast to this, Butt and Lahiri (2013) claim that light verbs have always been present in Indo-Aryan languages and propose a single underlying lexical entry for a light verb and its homophonous simple verb in order to account for their (co-)existence since the old Indo-Aryan stage. They also claim that light verbs are inert to any kind of change and hence, depart from the prevalent view that auxiliaries develop from further grammaticalization of light verbs.

Now, the existence of light verbs since old Indo-Aryan stage and their stable nature would imply that the number of light verbs present in different time periods of a language should roughly be the same. However, a pilot study of around a thousand sentence corpus of the following texts from different time periods of Marathi shows otherwise:

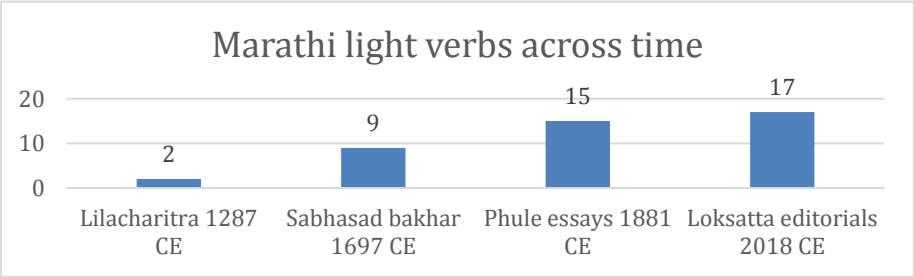


Figure 1: Marathi Light verbs across time

As seen in figure (1), number of light verbs attested in 13th century Marathi (two, in *Lilacharitra*) is significantly less than the number of light verbs attested in contemporary Marathi (seventeen, *Loksatta* editorials). This stark difference, along with the steady rise in the of number light verbs across time seen in this sample goes against the stability hypothesis put forth by Butt and Lahiri (2013), and necessitates a closer inspection of Marathi data in order to arrive at a better understanding of the diachronic nature of light verbs.
Therefore, by analyzing a thousand sentence corpus of various Marathi texts from time period ranging from 13th century C.E. to contemporary Marathi, the present paper shows that not all light verbs occur in all stages of Marathi, and there is a systematicity in the order of their emergence. Secondly, with the help of frequency patterns of the selected light verbs (*d͡zaɳe* (GO), *jeɳe* (COME), *gʰeɳe* (TAKE), *ɖeɳe* (GIVE), and *ʈakɳe* (THROW); selected based on their order of emergence and synchronic frequency), it tries to show that light verbs have undergone changes in terms of their collocational ability, frequency of occurrence, and to a smaller extent, their functions. Lastly, the paper takes into account Hook's (2001) tripartite classification of compound verb systems of Indo-Aryan languages (stage I being least advanced to stage III being most advanced in terms showing aspectogenesis) to show that Marathi, grouped under stage II, shows many characteristics of a stage III language; indicating that Marathi might slowly be moving towards a more advanced compound verb system like Hindi.

Thus, with the help of diachronic data, the present study traces the historical development of Marathi light verbs, and examines the implications of this evidence on our present day understanding of Indo-Aryan light verbs.

Selected References:

Beams, John. 1879. *Comparative Grammar of the Modern Indo Aryan Languages*. London: Trubner and Co.

Butt, Miriam & Lahiri, Aditi. 2013. Diachronic pertinacity of light verbs. *Lingua*. 135. 7–29.

Hook, Peter Edwin. 1991. The Emergence of Perfective Aspect in Indo-Aryan. In *Approaches to Grammaticalization II*, ed. Traugott, Elizabeth Cross and Heine, Bernd. 59-89. Amsterdam: John Benjamins.

Hook, Peter Edwin. 1993. Aspectogenesis and Compound Verb in Indo-Aryan. In *Complex Predicates in South Asian Languages,* ed. Verma, Mahindra. 97-113. New Delhi: Manohar Publishers and Distributors.

Hook, Peter Edwin. 2001. Where do compound verbs come from? (And where are they going?) In *Yearbook of South Asian Languages*. 101-130. De Gruyter Mouton.

Slade, Benjamin. 2013. The diachrony of light and auxiliary verbs in Indo-Aryan. In *Diachronica*. 30.4: 531–578.

Agreement in K'iche' (Mayan): Reflections on Microvariation

MARISA FRIED, PAULINA LYSKAWA, RODRIGO RANERO
UNIVERSITY OF MARYLAND, COLLEGE PARK

PURPOSE: We investigate microvariation in the realization of 3PL agreement in El Novillero K'iche' (Mayan), an under-described dialect of the language. We argue that a previously proposed agreement ban with inanimate arguments (England 2011) needs to be modified in order to account for microvariation and a broader data set. Based on targeted elicitation data, we show that for some speakers, 3PL inanimate agreement tracks the base-generated position (specifier vs. complement) of the agreement controller: arguments base-generated in a specifier position control agreement obligatorily, whereas arguments base-generated in a complement position optionally do so (Levin et al. 2019). Crucially, some speakers converge on this rule despite little evidence in the input. We then discuss microvariation among speakers, which we model as resulting from the differences in data distribution in the input; given variable acquisition histories, we predict that the kind of microvariation we observe would exist. The existence of this microvariation in this phenomenon also suggests the need for hypothesis-driven fieldwork (Davis et al. 2014) with individual speakers, instead of pooling data across speakers.

BACKGROUND: Variable agreement with 3PL arguments has been reported for a number of Mayan languages (England 2011 for a summary of the literature; Smith-Stark 1974 for Poqomam, Mateo-Toledo 2008 for Q'anjob'al, Zavala 1992 for Akateko, Aissen 1987 for Tzotzil, Henderson 2009 for Kaqchikel, Dayley 1985, Levin et al. 2019 for Tz'utujil). Each of these studies differ in the subset of constructions described and the methodology used for data collection. For example, England (2011) argues that K'iche' agreement is animacy-based: animates control agreement obligatorily, while inanimates do not control agreement at all. She bases her conclusion on a narrative corpus study; agreeing inanimates are unattested, as are non-agreeing animates:

(1)*K'iche corpus data distribution and pattern (England 2011: 402)*

	TRANSITIVE SUBJECT	TRANSITIVE OBJECT	INTRANSITIVE SUBJECT
ANIMATE	common & all agree	rare & all agree	common & all agree
INANIMATE	unattested	common & none agree	rare & do not agree

NOVEL DATA: Targeted elicitation data from K'iche show that England's account is incomplete. As the data below show, a 3PL absolutive (A3PL) inanimate object agrees *optionally*:

(2) Oj x-**e**-/**ø**-qa-q'aluj **e-keb'** **che'**.
1PL CPL-**A3PL**-/**Ø**-E1PL-hug **PL-two** **tree**
'We hugged two trees.'

More strikingly, A3PL inanimates agree *obligatorily* in other constructions, e.g. positional predicates in (3). Positional predicates are a distinct class of unergative one-place predicates in Mayan that generally have stative-like meanings (Henderson 2019, Lyskawa & Ranero 2019).

(3) **E**-/***ø**-tzalan **keb'** **ab'aj** cho le xan.
A3PL/***ø**-lean **two** **stone** against the wall
'Two stones lean against the wall.'

We observe obligatory agreement with inanimate 3PL ergative (E3PL) transitives subjects as well:

(4) *Context*: a windy storm passed by the night before, causing two trees to fall through your door.
Le **keb'** **che'** x-ø-**ki**-/***u**-/***ø**-wulij le uchija.
The **two** **tree** CPL-A3PL-**E3PL**-/**E3SG**-/**ø**-break the door
'The two trees broke the door.'

We summarize the novel pattern obtained via elicitation with a single speaker below, highlighting the cells that England (2011) does not discuss explicitly in her corpus study:

(5) *3PL inanimate agreement in K'iche; consultant A*

TRANSITIVE SUBJECT	TRANSITIVE OBJECT	POSITIONAL SUBJECT	UNACCUSATIVE SUBJECT
obligatory	optional	obligatory	optional

The puzzle is the following: even though certain data types with 3PL inanimate subjects are very infrequent, speakers can acquire a grammar displaying obligatory agreement in such data. How can we model the acquisition path leading to this final grammar?

ANALYSIS: To recapitulate, England finds no tokens in her corpus study of inanimate subjects of transitives (4), nor does she single out any instances of inanimate subjects of positional predicates (3). The distribution of the corpus data in (1) is consistent with an analysis of animacy-based agreement. However, when probing for the rare tokens directly via targeted elicitation, our K'iche consultant provides systematic judgments that refute the animacy-based hypothesis. As a result, we propose the following: acquirers start out with the hypothesis that *all* arguments must agree. A speaker will only modify this hypothesis if they encounter positive evidence of non-agreeing arguments. Given the paucity of data in the cells highlighted above, some speakers never modify their initial hypothesis. In other words, they are so rarely exposed to these data points that they maintain the hypothesis that inanimates must control agreement. In contrast, other speakers *are* exposed to enough tokens of non-agreeing inanimates to modify their hypothesis and settle on optional agreement. If they are exposed to non-agreeing inanimates either in a transitive object or intransitive subject position, they hypothesize that optionality is allowed only for arguments base-generated as complements. This kind of analysis has proven fruitful in modelling 3PL agreement in closely related Tz'utujil (Levin et al. 2019). On the other hand, if the speakers are exposed to at least some instances of non-agreeing inanimate transitive or positional subjects, they hypothesize that optionality holds for all constructions. Our proposal predicts the kind of microvariation we might observe across speakers of a single dialect. As evidence, consider the data from consultant B below, who displays optionality across the board with inanimate arguments:

(6) *3PL inanimate agreement in K'iche; consultant B*

TRANSITIVE SUBJECT	TRANSITIVE OBJECT	POSITIONAL SUBJECT	UNACCUSATIVE SUBJECT
optional	optional	**optional**	optional

We assume that consultant B's acquisition history was different than A's: whereas speaker A was not exposed to enough tokens of certain non-agreeing inanimates to modify her initial hypothesis, speaker B was. As a result, speaker B modified her hypothesis in all cells, arriving at a grammar which displays optional agreement with inanimates regardless of structural position. In our talk we will also discuss agreement with animate controllers which shows a parallel pattern.

CONCLUSION: Whereas agreement variability with 3PL arguments has been discussed before in Mayan, we coincide with Levin et al. (2019) in noting that microvariation can shed light on different acquisition trajectories. England 2011's corpus investigation gives us a glimpse into the data distribution that speakers might be exposed to in the acquisition process, thus allowing us to model the different grammars that speakers settle on, in particular for those data points that are extremely rare in the input.

SELECTED REFERENCES: England, N. 2011. Plurality Agreement in Some Eastern Mayan Languages. *International Journal of American Linguistics*. **Davis et al. 2014**. How to investigate linguistic diversity. *Language*. **Levin, T., Lyskawa, P. & R. Ranero. 2019.** Optional agreement in Santiago Tz'utujil is syntactic. *41st Annual Conference of the German Linguistic Society*. **Lyskawa, P. & R. Ranero**. Optional agreement in Santiago Tz'utujil. Ms, University of Maryland, College Park.

Anaphora Resolution of Mandarin Reflexive *Ziji*: An ERP Study

Xu YAN, Yuxia WANG
School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China, 200240

Anaphora resolution is to establish the antecedent-pronoun dependency, a core issue in natural language processing. English reflexives are strictly bound in its local domain according to Binding Principle A (Chomsky, 1981); Mandarin reflexive *ziji* does not follow the principle strictly. Hu and Pan (2002) argued that *ziji* could refer to any prominent antecedents, when there was no intervening noun to block the referential process. In the past 20 years, many ERP studies were conducted to figure out the neural activities of anaphora resolution, yet few of them were in Chinese. Among the few, Li and Zhou (2010) investigated the binding preferences of Mandarin reflexive *ziji* and reported the effect of P300 and P600 for the long-distance binding condition relative to the local/ambiguous binding conditions. Interestingly, Dillon (2014) found that not all locality effects may be reduced to the effects of temporal decay and retrieval interference. Given the scarcity of and the controversies among the existing literature, the current study aimed to examine how binding distance and animacy affect Mandarin reflexive *ziji* using ERPs.

We recruited 24 native Mandarin speakers and adopted a 2*2 within-subject designs via manipulating binding distance (local and distant binding) and animacy feature of antecedents (single and double animate). The experiment included 120 target sentences and 140 fillers in a pseudo-random sequence with a Latin-square design. Prototype sentences were given below.

1a) DoubleAminate_LocalBinding (DistantBiding)

病人（医生）/ 说 / 医生（病人）/ 担心 / 自己 / 明天 / 主刀的 / 心脏 / 手术。

Patient (doctor) / said / doctor (patient) / worried about / *ziji* / tomorrow / operate DE/ cardiac / surgery.

1b) SingleAminate_LocalBinding (DistantBiding)

医院（医生） / 说 / 医生 （医院） / 担心 / 自己 / 明天 / 主刀的 / 心脏 / 手术。

Hospital (doctor) / said / doctor (hospital) / worried about / *ziji* / tomorrow / operate DE/ cardiac / surgery.

The target sentences (as mentioned above) have two NPs as potential antecedents, and the disambiguity verb is semantic congruent with one of the profession nouns. If *ziji*-preferred antecedent retrieved at the segment *ziji* is congruent with the semantic requirement of the verb, the processing at the verb region is less effortful; otherwise, it is more difficult. Moreover, animacy might also play a certain role. Hence, we predicted that the local binding and single animate condition could bring less processing cost.

Our results reported a reliable N400 effect ($p<.1$) in 400-430ms with the double animate, local binding condition (in red) relative to the double animate, distant binding condition (in green) taking CP1 as representative (Figure 1a & 1b). A significant P600 effect ($p<.05$) in 580-650ms was shown with the single animate, local binding condition (in red) compared with the single animate, distant binding condition (in green) taking CZ as a representative (Figure 2a & 2b). The P600 effect happens most in the left hemisphere, and the dorsal part of the right. Focusing on animacy, the results illustrated a prominent N400 ($p<.01$) and some P600 effect ($p<.05$) in 400-700ms between the double animate, local binding condition (in red) and the single animate, local binding condition (in green) showed in Figure 3a & 3b. The double animate, distant binding condition (in red) and single animate, distant binding condition (in green), demonstrated a P600 effect ($p<.05$) in 600-650ms which mainly centered on the left central area (Figure 4a & 4b).

Our results suggested that people preferred the distant binding than the local binding, which is in violation with Chomsky's Principle A, especially when the reflexive is preceded by one animate antecedent.. Our results were consistent with the information sensitivity theory (Kaiser, 2009) and first-mention effect (Murdock, 1962). Meanwhile, the double animate antecedents triggered stronger potentials compared with the single animate antecedent, which agreed with most studies (Dillon, 2014): even in the local binding condition, subjects still held uncertainty of and needed to reanalyze their choice, as both N400 and P600 effect were associated with double animate conditions. In sum, both binding distance and animacy affected the resolution of Mandarin reflexive *ziji*.

References: [1] Brian, D. et al. (2014). The structure-sensitivity of memory access: evidence from mandarin chinese. *Frontiers in Psychology*. [2] Li, X. , & Zhou, X. . (2010). Who is *ziji*? ERP responses to the chinese reflexive pronoun during sentence comprehension. *Brain Research.* [3] Lu, D.& Garnsey， S. (2008). Whose self do you mean?—Processing Mandarin Chinese reflexives. Poster presented at *the 21th CUNY*

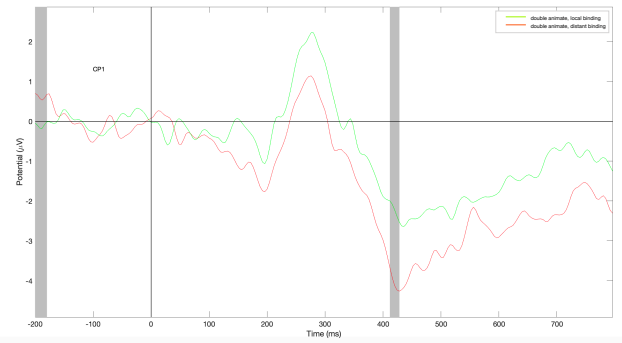


Figure1a. p < 0.05

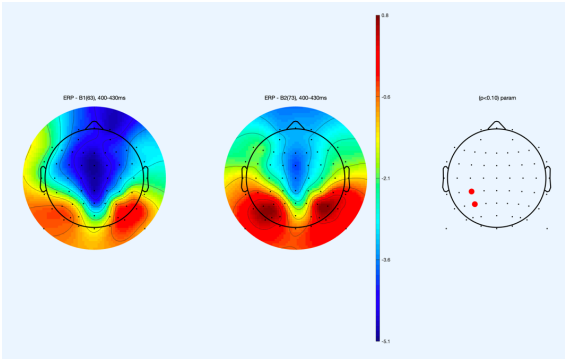


Figure1b. The left, the local; the right, the distant

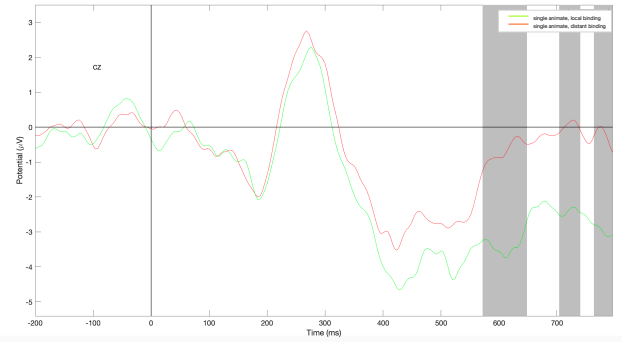


Figure2a. p < 0.05

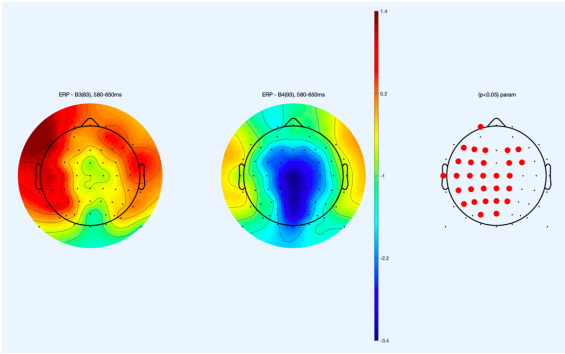


Figure2b. The left, the local; the right, the distant

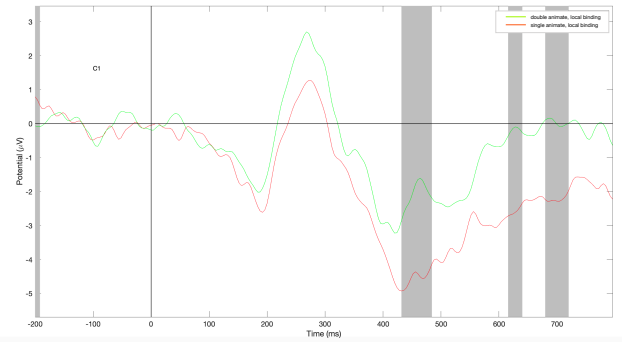


Figure3a. p < 0.01

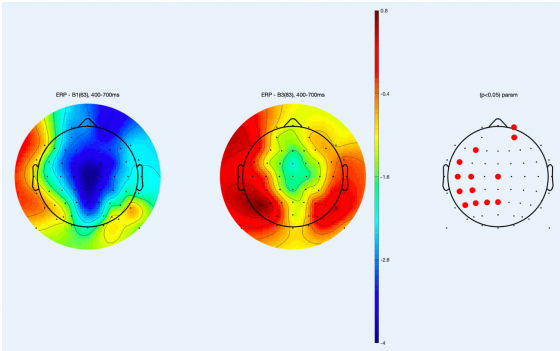


Figure3b. The left, the double; the right, the single

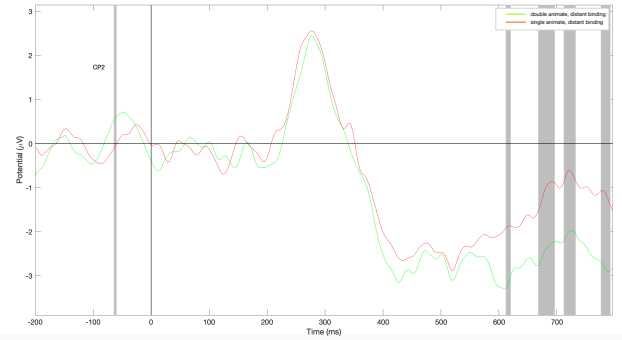


Figure4a. p < 0.05

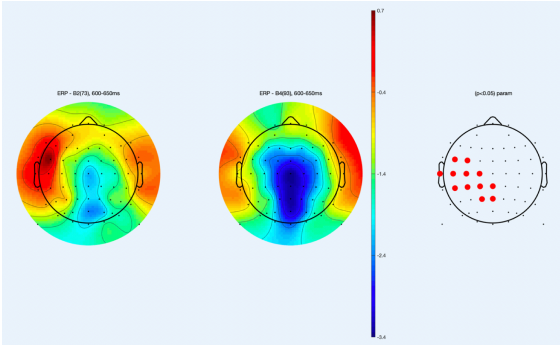


Figure4b. The left, the double; the right, the single

Dative Causees in Basque

Joe Class
Yale University

Introduction Basque exhibits substantial dialectal variation in its case assignment patterns for causees. When the subjects of unergative verbs are embedded within a causative, they are assigned the absolutive case in northern Basque, as opposed to the dative case in southwestern Basque. This has substantial implications for competing theories of case assignment for causees, as northern Basque provides support for Arregi 2018 and Baker 2015's dependent case analysis, whereas southwestern Basque fits well within Ippolito 2000's applicative analysis. In this paper, I will argue that causees in all Basque dialects receive case via dependent-case marking, and that an underspecified spell-out rule between the absolutive and dative case in southwestern Basque can account for dative surface forms.

Theoretical Framework In all Basque dialects, causees in embedded transitives receive dative case, as in the following.

(1) *Ama-k* ***haurr-ari*** *zopa* *jan-erazi* *dio* (Donohue 2007:127)
Mother-ERG **child-DAT** soup.ABS eat-CAUS has.3DATS
'The mother made the child eat the soup'

Both applicative models and dependent case frameworks for causatives can account for embedded transitives, but not for the dialectal variation in embedded unergatives. In discussing these theories, I employ Harley 2017's model of causatives as a basic syntactic framework. In this model, the causer is introduced in the specifier of VoiceP, and the Voice head selects for a CausP which governs lower verbal structure.

Applicative Causees Ippolito 2000 proposes that subjects embedded within a causative don't constitute external arguments of the embedded verb, but rather applied arguments. Syntactically, this is compatible with Kratzer 1996's work arguing that verbs don't select for external arguments themselves, but rather that they are optionally introduced by a Voice head. As an applied argument, the causee is introduced in the specifier of an ApplP (as in Pylkkanen 2000), and receives dative case from the Appl head through spec-head agreement. In our model, this would specifically involve a Caus head in CausP selecting for an ApplP, which itself c-commands a lower verbal projection (vP). This theory can account for embedded transitives in all of Basque, as the causee simply receives dative from the Appl head. Looking at unergatives, data from southwestern Basque also supports this analysis, as all causees that are the subjects of unergative verbs receive the dative case.

(2) *Norbait-ek* *eta* *zerbait-ek* *iraun-arazi* *dio* ***hizkuntz-ari*** (Rijk 2008:380)
Someone.ERG and something.ERG last-CAUSE has.3DATS **language-DAT**
'Someone and something has caused the language to last'

Dependent Case-Marking Dependent case-marking, first formulated by Marantz 1991, claims that certain cases are assigned to a DP as a reaction to its structural relation to other DPs in a sentence or phrase. Looking at causatives, Arregi 2018 and Baker 2015 propose that the verbal structure below an external argument-introducing VoiceP (analgous to CausP in my model) constitutes its own domain for dependent case. Specifically, if one DP in a CausP c-commands another lower DP, it receives the structural dative case. For Basque, this theory predicts that causees of embedded transitives will receive the dative case, but causees of embedded unergatives will receive the default absolutive case. In northern Basque, causees of unergative verbs do receive the absolutive case, providing support for a dependent case-marking analysis.

(3) *Errege-ak* ***gerla*** *iraun-arazi* *zuen* (Oyharcabal 2003:231)
King-ERG **war.ABS** last-CAUS has
'The king caused the war to last'

Initially, the opposing patterns from northern and southwestern Basque seem to suggest that they employ different mechanisms for assigning case to causees. There is promising evidence, however, that unergative causees in southwestern Basque are truly absolutive, receiving dative morphology as the result of an underspecified spell-out rule. This supports a dependent case-marking analysis for all of Basque.

DOM in Southwestern Basque Odria 2014 points out that southwestern Basque exhibits a widespread pattern of Differential Object Marking (DOM), where definite, animate DPs receive the dative case in configurations that usually license the absolutive case, such as the direct object.

(4) *Ni-k zu-ri ikusi di-zu-t* (Odria 2014:291)
I-ERG you-DAT see have-2DATS-1ERGS
'I have seen you'

Odria provides evidence supporting the claim that DOM objects in Basque are truly absolutive, basing her analysis on several structural properties (beyond basic position) that they share with absolutive objects. For example, DOM objects can license descriptive secondary predication (6), unlike dative indirect objects, which are commonly analyzed as low applied arguments (5). Like DOM objects, **dative causees of embedded unergatives** can also license secondary predication (7).

(5) *Ni-k*$_i$ ***amona-ri***$_j$ *umea*$_k$ ***pozik***$_{i/*j/k}$ *eraman di-o-t* (Odria 2014:294)
I-ERG **grandmother-DAT** child.ABS **happy** carry have-3DATS-1ERGS
'I carried the child to grandmother, happy.'

(6) *Nik* ***zu-ri***$_i$ ***pozik***$_i$ *ikusi di-zu-t* (Odria 2014:295)
I.ERG **you-DAT** **happy** see have-2DATS-1ERGS
'I have seen you happy'

(7) *Jonek* ***Miren-i***$_i$ *trampolin-etik* ***mozkortuta***$_i$ *salt-arazi dio* (Odria 2014:296)
John.ERG **Mary-DAT** trampoline-ELA **drunk** jump-CAUS have-3DATS
'John made Mary jump from the trampoline drunk'

Here it can be seen that dative causees in southwestern Basque don't behave like applied arguments, but rather like DPs in a position that canonically licenses the absolutive case. Given that causees of unergatives are absolutive in the northern dialects, I propose that they are also syntactically absolutive in the southwestern dialects. In contrast to the northern dialects, they appear on the surface with dative morphology as a result of DOM.

To account for these dative surface forms, I propose that southwestern Basque exhibits the same surface morphology between the absolutive and dative cases in certain contexts. Specifically, southwestern Basque has an underspecified spell out rule for definite objects, animate objects, and causees, by which the same dative morphology is used for these DPs when merged in a position that licenses the dative or absolutive case. This is similar to Barany 2018's proposal for DOM phenomena in Spanish and Hindi, whereby underspecified spell out rules between the dative and the accusative give rise to case syncretism.

Dependent Case-Marking for all Basque Dialects Since unergative causees in all Basque dialects are syntactically absolutive, I conclude that dependent case-marking is the mechanism by which causees receive case in Basque. When transitives are embedded, the causee in CausP c-commands a lower DP, and receives structural dative case. When unergatives are embedded, the causee doesn't c-command any lower DP, and receives the default absolutive case. In the northern dialects, this DP appears on the surface as absolutive. In the southwestern dialects, an underspecified spell out rule between dative and absolutive for causees results in dative surface morphology. In combination with Harley's analyses of Japanese and Italian causatives (1995, 2007), as well as Baker 2015's general theory of case assignment in causative constructions, this conclusion lends support to a unified account of dependent case-marking for causatives across all languages.

Variation in subject-verb agreement in the history of Scots

Lisa Gotthard, University of Edinburgh

This paper investigates the interaction between the *Northern Subject Rule* and *do*-support in the period when *do*-support emerged in Scots. Scots is a variety closely related to English, which is historically spoken in lowland Scotland. The history and development of Scots in relation to English can be described through its processes of divergence from a shared ancestor, starting in the 11th century, and later, in the 16th-18th century, convergence with English again during the period of *anglicisation* of Scots; in this period, social and political shifts between Scotland and England caused Scots to fall under more influence from (Southern) English, causing distinctively Scots features to decline in favour of English ones. Comparative-syntactic studies on Scots and English can thus provide insight into the syntactic outcomes of long-term contact between closely related varieties, during which time their relative socio-political prestige shifts.

The Northern Subject Rule (NSR) is a present tense inflectional system which has historically been operational in Northern English, Irish English, and Scots, and versions of it still operates variably in these varieties today. The NSR requires structural adjacency between plural pronoun subjects and the finite verb in order to establish subject-verb (s-v) agreement; as analysed by de Haas (2011), if the adjacency or subject type requirements are not met, the verb takes a *default inflection*, e.g. -(i)s in Scots (1a-b), whereas if subject-verb agreement is successful, the verb has ∅-inflection (1c; examples 1a-c fabricated for purpose of demonstration of the rule).

(1) a. He/she/the bird(s) sing-is
b. We/you/they quietly sing-is
c. We/you/they sing-∅

Thus, the NSR can be analysed as a system equal to *do*-support; the default inflection in an NSR system is a last-resort operation when s-v agreement fails, like the insertion of the auxiliary *do* in a *do*-support system. From this, we can postulate that when *do*-support emerged in Scots during the anglicisation period, it entered grammar competition not only with main verb movement, but also with the NSR – both still productive systems in Scots at the time, not in competition with each other but operating simultaneously. This idea of *do*-support entering competition on two fronts finds some support by trends reported in the literature (Jonas 2002; Meurman-Solin 1993; Gotthard 2019) that Scots *do*-support emerged later and at a slower rate than in English. Therefore, this paper investigates whether the introduction of *do*-support in Scots was subject to NSR constraints, and hypothesises that *do*-support would be adopted first in environments where NSR default inflection would apply, i.e. environments which requires last-resort operations to establish s-v agreement, and, crucially, *do*-support would be resisted in environments where the subject type and adjacency requirements of the NSR are establishing successful s-v agreement. This hypothesis finds some support from present-day variation: in Buckie Scots, a present-day Scots variety, Smith (2000) finds optional use of *do*-support in negative declaratives, but only with the NSR subject types (i.e. plural pronouns; see (2), from Smith 2000: 232).

(2) a. They ∅ na seem to bide in the Beacons lang
They ∅ NEG seem to stay in the Beacons a long time
b. They dinna ken they're gan to wear a kilt.
They DO+NEG know they're going to wear a kilt

This Buckie pattern may indicate an NSR subject constraint on the development of *do*-support in Scots, which the present paper quantitatively evaluates by using a part-of-speech tagged version of the *Helsinki Corpus of Scottish Correspondence* (CSC, Meurman-Solin and VARIENG). The CSC consists of over 400,000 words of correspondence data, produced between 1540 and 1750 – the period of anglicisation of Scots. The corpus is being enriched with syntactic annotation in the *Penn Parsed Corpora of Historical English* format (e.g. Kroch and Taylor 2000; Kroch et al. 2004) as part of my PhD project, and the parsed version will be made available

upon completion; the new parsed CSC will allow us to investigate the presence of these s-v agreement strategies in 16th-18th century Scots, by collecting quantitative results which we can more readily compare to similar studies made on English (e.g. Kroch 1989; Ellegård 1953).

Preliminary findings from the data up to 1650 indicate that there is indeed a dispreference for negative declarative *do*-support with plural pronoun subjects after 1600 (there is only 1 example (3%) of *do*-support before 1600 (3a)), whereas *do*-support dominates with all other subject types. However, the results also indicate differences between gender groups: female writers predominantly use *do*-support with all subjects, including plural pronouns (3b), but male writers exhibit more variability with respect to subject type; while male writers prefer verb-raising with plural pronouns (3c), as the hypothesis predicts, they also use more *do*-support than verb-raising with 3sg pronouns (3d-e), which is not predicted by the hypothesis.

(3)	a.	for I dow not get yam(='them') to hym	(male, 1548)
	b.	Zie(='you') did not wische me oftiner nor(='than') I did my self	(female, 1617)
	c.	othewise[sic] they durst not so rune away without ordour	(male, 1646)
	d.	If It doe not please god	(male, 1635)
	e.	Iff he be not ffund thair	(male, 1616)

These preliminary results will be substantiated with the addition of the remainder of the data from the CSC.

References

Ellegård, A. (1953). *The auxiliary do*. Stockholm: Almqvist and Wiksell.

Gotthard, L. (2019). Why *do*-support in scots is different. *English Studies* 100(3), 1-25.

de Haas, N. (2011). Morphosyntactic variation in Northern English. Ph. D. thesis, Utrecht: LOT.

Jonas, D. (2002). Residual V-to-I. In *Syntactic Effects of Morphological Change*, 251-270. Oxford: OUP.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *LVC* 1, 199-244.

Kroch, A., B. Santorini, and L. Delfs (2004). The Penn-Helsinki Parsed Corpus of Early Modern English, release 3. UPenn. [https://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE- 3/index.html].

Kroch, A. and A. Taylor (2000). The Penn-Helsinki Parsed Corpus of Middle English, 2nd edition, release 4. UPenn. [https://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE- 4/index.html].

Meurman-Solin, A. and VARIENG (Accessed 2018). Helsinki Corpus of Scottish Correspondence (1540-1750). University of Helsinki. [http://www.helsinki.fi/varieng/csc/manual/].

Meurman-Solin, A. (1993) Periphrastic Do in Sixteenth- and Seventeenth Century Scots. In *Variation and Change in Early Scottish Prose*, 259?276. Helsinki: Suomalainen Tiedeakatemia.

Smith, J. (2000). Negative *do* in Buckie Scots. *English World Wide* 21(2), 231-259.

Body partonomy as evidence of language shift: Emberá senses of ‘hand’ and ‘foot’ in the Spanish of western Colombia

Eliot Raynor
Indiana University, Bloomington
Princeton University

Monolingual varieties of Spanish in the Colombian provinces of Chocó, Antioquia, and Córdoba demonstrate ambiguity in the referential senses of *mano* ‘hand’ and *pie* ‘foot’, as shown in the examples below drawn from the PRESEEA (2014) corpus for Medellin:

(1) No podía mover las **manos** porque me quebré por acá.
NEG could move DEF hands because DAT-1SG broke around here
‘I couldn’t move my **hands/arms** because I had a break around here.’ (MEDE_MO2_2)

(2) Se me partió la carne del **pie** izquierdo al lado de la rodilla.
REFL DAT-1SG opened DEF flesh of-DEF foot left to.the side of DEF knee
‘The skin of my left **foot/leg** was cut open on the side of my knee.’ (MEDE_H03_5)

Cross-linguistic surveys have revealed that colexifications of ‘hand’-‘arm’ and ‘foot’-‘leg’ are rather common outside of Western Europe (Brown, 2013; Koptjevskaja-Tamm & Liljegren, 2017; List et al., 2018). This pattern is also frequently found in contact varieties for which one or more substrate languages exhibit similar semantic overlap—as discussed, for instance, in Cassidy and LePage’s (1980) *Dictionary of Jamaican English* and Holm’s (1978) monograph on Miskito Coast Creole (p. 217); for a broad survey, see Parkvall and Baker (2012, pp. 237-239). Within the Romance language family, ‘hand’-‘arm’ and ‘foot’-‘leg’ polysemy appear to have emerged only in situations of intense language contact, such as was the case for Romanian (McLure, 1975) as well as for Portuguese-lexified creoles such as Saotomense (Ferraz, 1979, p. 100-101).

Noting that the use of Spanish *mano* and *pie* in reference to the entire limb has not been reported in prior work on Spanish variation, the presence of this pattern in monolingual varieties of Spanish across western Colombia motivates two research questions that guide the present study:

(i) How acceptable are *mano* for ‘arm’ and *pie* for ‘leg’ by Colombian Spanish speakers?

(ii) With what degree of certainty can we determine whether language-internal mechanisms (i.e., metonymy) versus contact-induced ones (i.e., substrate transfer) account for the development of this usage across western Colombia? If this comprises a contact-induced change, which substrate language(s) were involved?

In response to (i), an acceptability judgment task designed in Qualtrics and completed by monolingual Spanish speakers (n=83) demonstrates categorical acceptance of *mano* and *pie* referring to ‘arm’ and ‘leg’, respectively, by speakers from the Colombian provinces of Antioquia and Córdoba. With respect to (ii), the present study also provides historical and linguistic evidence that the semantic ambiguity of *mano* and *pie* in Colombian Spanish most plausibly derives from language contact than from internal mechanisms of change. The historical timeline presented here focuses on the western frontier of New Granada (Colombia) in the 17th century, a period during which indigenous communities were forced into sedentary *reducciones* (Williams, 2004). During this period speakers of Amerindian languages of the Chocoan family shifted en masse to Spanish and thus introduced to the Spanish varieties of the region the semantics of body partonomy from Emberá/Wounaan, in which ‘hand’ and ‘arm’ and ‘foot’ and ‘leg’ are not distinguished—e.g., in Emberá Chamí *húa* refers to ‘upper extremity’ or ‘arm including hand’ and *hĩ́rũ* ‘lower extremity’ or ‘leg including foot’ (Huber & Reed, 1992, pp. 19, 29; Loewen, 1957, pp. 44a, 219a).

References

Brown, C. H. (2013). Hand and Arm. In Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/129.

Cassidy, F. G., & Le Page, R. B. (1980). *Dictionary of Jamaican English* (2nd ed.). Cambridge: Cambridge University Press.

Ferraz, L. I. (1979). *The creole of São Tomé*. Johannesburg: Witwatersrand University Press.

Holm. (1978). *The creole English of Nicaragua's Miskito Coast: Its sociolinguistic history and a comparative study of its lexicon and syntax* (Ph.D. dissertation, University of London, University College).

Huber, R. Q., and Reed, R. B. (1992). *Vocabulario comparativo: Palabras selectas de lenguas indígenas de Colombia*. Bogotá: Asociación Instituto Lingüístico de Verano.

Koptjevskaja-Tamm, M., & Liljegren, H. (2017). Semantic patterns from an areal perspective. In R. Hickey (Ed.), *The Cambridge handbook of areal linguistics* (pp. 204-236). Cambridge: Cambridge University Press.

List, J., Greenhill, S., Anderson, C., Mayer, T., Tresoldi, T., & Forkel, R. (Eds.). (2018). *Database of Cross-Linguistic Colexifications*. Jena: Max Planck Institute for the Science of Human History. http://clics.clld.org.

Parkvall, M., & Baker, P. (2012). Idiomatic (potential) calques and semantic borrowing. In Bartens, Angela & Baker, Philip (eds.), *Black through white: African words and calques which survived slavery in creoles and transplanted European languages* (pp. 231-248). London: Battlebridge Publications.

Loewen, J. (1957). *Comparative dictionary Choco dialects-Spanish-English.* Cali, Colombia: Mennonite Brethren Board of Missions.

McClure, Erica F. (1975). Ethno-anatomy: The structure of the domain. *Anthropological Linguistics*, *17*, 78-88.

PRESEEA. (2014). *Corpus del proyecto para el estudio sociolingüístico del español de España y de América*. Alcalá de Henares: Universidad de Alcalá. http://preseea.linguas.net.

Williams, C. (2004). *Between Resistance and Adaptation: Indigenous Peoples and the Colonisation of the Chocó, 1510-1753*. Liverpool: Liverpool University Press.

Stress shift accompanying verb suffixation in Gujarati

Chenchen Wang
University of California, Los Angeles

Gujarati is an Indo-Aryan language primarily spoken in the state of Gujarat in northwest India. It has approximately 45 million speakers as of the 2001 census. Stress in Gujarati has attracted much interest among linguists. Gujarati is a synthetic agglutinative language, which has a systematic and extensive suffixation system. Suffixes can be applied one after the other, with new suffixes affixing to previous ones. This paper examines stress shift accompanying suffixation in Gujarati causatives and passives, and shows that in at least causatives and passives, Gujarati maintains a well-defined stress pattern.

Gujarati causatives and passives are derived by suffixation to the verb roots, which can be further suffixed to derive additional verb forms such as second causatives (Cardona 1965:114-118). Gujarati causative suffixes are -ɑʋ, -ɑɖ, -ɖɑʋ, -ʋ, and -eɖ. The Gujarati passive suffix is -*a*. In causatives and passives derived by suffixation, three types of root alternations can occur (Cardona 1965:112): (1) /ʋ/ epenthesis: ʋ, which has two allophones [w] and [v], is inserted before the suffix if the word ends in a vowel or [ɦ]; (2) /ɑ/ reduction: the [ɑ] in the last syllable of the root is replaced by a [ə] if the suffix has an [ɑ]; and (3) /ə/ deletion: [ə] of the final syllable of the root is deleted if the suffix starts with a vowel.

To investigate stress shift in Gujarati, data were collected from a consultant who was a third year undergraduate student. He is a fluent heritage speaker of standard Gujarati. He was born in the U.S. to parents who speak only Gujarati. He moved to Gujarat with his family when he was thirteen, where he lived and attended school until coming back to the U.S. for college.

Representative data are shown below:

(1) [ˈkʰɑ] ("eat") vs. [kʰəwˈɖɑw] ("feed") (ʋ-epenthesis)
(2) [ˈkɑp] ("cut") vs. [kəˈpɑw] ("cause to cut") vs. [kəˈpɑwɖɑw] ("cause to cause to cut") vs. [kəˈpɑ]("be cut") (ɑ-reduction)
(3) [səˈmədʒ] ("understand") vs. [səmˈdʒɑw] ("explain") (ə-deletion)

Root alternations and stress were analyzed using Praat. Exemplary data and analysis are shown in Fig.1 and Tables 1 and 2. Fig.1 shows the spectrogram of [kəˈpɑwɖɑw] ("cause to cause to cut") in (2). Table 1 shows F1 (Hz) and F2 (Hz) of words in (2). Table 2 shows parameters relating to stress, i.e., intensity, pitch, and vowel length, (Ladefoged 2011: 111) for words in (2). The spectrogram such as the one shown in Fig.1 for each word was used to determine the parameters such as those listed in Tables 1 and 2, which were then used to generate the transcriptions of the words. From Table 2 it can be seen that the stressed syllable exhibits higher intensity and longer vowel length, but not higher pitch.

The results show that at least in the context of suffixation in causatives and passives, Gujarati, by shifting the stress to the suffixes, maintains a well-defined stress pattern, i.e., the ultimate syllable in di-syllabic words and the penultimate syllable in tri-syllabic words. The corresponding root alternations are consistent with this observation because /ɑ/ reduction is a well-known process to de-stress a syllable (Ladefoged 2011: 109-110 and Crosswhite:1), while /ə/ deletion is a common phonological process in Indo-Aryan languages, which appears to function as removing a non-stress syllable.

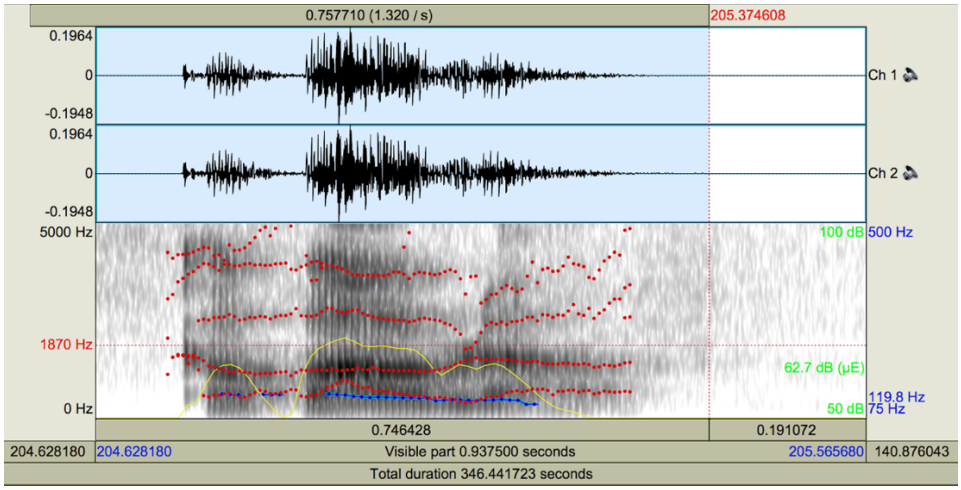


Fig. 1 Spectrograms of [kəˈpʊɖʊ] (‘cause to cause to cut’) in (2).

Table 1. F1 (Hz) and F2 (Hz) of words in (2)

	[ˈkɑp]	[kəˈpɑw]	[kəˈpɑwɖɑw]	[kəˈpɑ]
1st syllable	913 1438	568 1308	667 1231	660 1265
2nd syllable		697 1248	785 1270	664 1268
3rd syllable			747 1427	

Table 2. Intensity, pitch, and vowel length of words in (2)

	[ˈkɑp]	[kəˈpɑw]		[kəˈpɑwɖɑw]			[kəˈpɑ]	
		1st syl	2nd syl	1st syl	2nd syl	3rd syl	1st syl	2nd syl
Intensity (dB)	61	62	63	64	68	55	65	67
Pitch (Hz)	undefined	120	113	128	121	undef.	129	124
Vowel length (sec)	0.19	0.08	0.20	0.08	0.15	0.13	0.10	0.2

References

Cardona, George. *A Gujarati Reference Grammar*. Philadelphia: University of Pennsylvania Press, 1965.
Crosswhite, K. (2018, June 3). *Vowel Reduction*. Reading presented at Seminar in University of Chicago, Chicago. Retrieved from http://washo.uchicago.edu/seminar/crosswhite-reduction.pdf
Ladefoged, P. and Johnson, K., A Course in Phonetics. 6th ed., 2011.

Manual and Nonmanual Negation in American Sign Language: A Corpus Study

Introduction Negation is one of the most descriptively well-studied phenomena in sign languages, especially the fact that sign languages have both "manual" signs expressing negation (eg. *NOT, NONE in ASL)* as well as "nonmanual" markers (e.g. head-shaking). However, research on negation in sign languages has primarily been based on elicited judgments from sentences that are either created by starting from a spoken language equivalent, or are an in-depth focus of a small number of manual signs based on introspection. In this study, we investigate naturally-occurring expressions of negation in a corpus of publicly available videos of American Sign Language (ASL) to ask how previous observations about manual signs in ASL extend to a wider range of expressions of negation, and to better understand the relationship between manual and nonmanual expressions of negation.

Source We analyzed 10 videos from public social media posts, ranging from 2 min. 30 seconds to 25 min. (Table 1), for a total of 97 minutes of signing. Videos were chosen for featuring a range of ages, genders, and contexts; all featured Deaf signers of ASL (8/10 known to be native signers, other 2 unknown age of acquisition) signing directly toward the video camera. Each video was coded for expressions of negation by one of the authors, a Deaf signer of ASL. The result was a total of 125 tokens of negation expressed, via 12 different types (Table 2).

Results Our corpus study supported some previous findings from the most comprehensive introspection study of ASL negation (Wood 1999), containing productions of all of the four signs she studied in depth: NOT, NO^{O}, NEVER, and NOTHING. The most frequent negation type seen in our corpus was in fact the sentential negation sign NOT. However, our data also showed some surprising divergences that were not predicted by existing research on negation in sign languages. The first of these was, notably, the high frequency of the headshaking *nonmanual* marker used as the sole negative marking with its own timeslot: negative headshake appeared simultaneously with a negative sign 8 times, but appeared as the *sole marker of negation* another 25 times, the second most frequent negation type in the corpus. The second unexpected finding from this corpus was that negative nonmanual headshaking and negative manual signs show a double dissociation: there are 29 examples of negative headshake occurring without any negative manual sign (1) and more surprisingly 25 examples occurring without any manual sign at all (e.g. the headshake has its own timeslot) (2), while there are 43 examples of negative manual signs occurring without negative nonmanual marking (3).

Analysis Our results suggest some discrepancy between judgements about negation in ASL based, on the one hand, traditional elicitation and introspection work, and on the other hand the results of a corpus study of naturally occurring negation. In particular, there were two structures that have been reported to be ungrammatical in ASL but which occurred multiple times, across multiple signers, in our corpus: (a) the use of headshake nonmanual as a stand-alone negative marker with its own timeslot and (b) the use of negative manual signs in ASL without nonmanual headshake marking. This latter also raises doubts about one of the most well-known generalizations about ASL as a "nonmanual dominant language", since there were several clear examples where negative nonmanual marking is not necessary (cf. Zeshan 2006).

Table 1.

Video	M/F	Age	Native Signer (Y/N)	Video Length
A	M	33	Y	3 min. 8 sec.
B	F	Late 30s	Unknown	8 min. 10 sec.
C	F	30s	Y	3 min. 21 sec.
D	F	66	Y	6 min. 56 sec.
E	M	~13, 14	Unknown	3 min.
F	M	30	Y	14 min. 28 sec.
G:	F	50s	Y	17 min. 12 sec.
H	F	Late 30s	Y	12 min. 46 sec.
I	F	Early 30s	Y	2 min. 33 sec.
J:	F	Late 40s	Y	25 min. 27 sec.

Table 2.

Negative Expressions	Videos	Token Count
***NOT**	A, B, D, G, H, I, J	27
Negative headshake only	A, D, E, G, J	25
CAN'T	D, F ,G, I, J	17
NO (palm outward)	B, C, G, H, J	13
NO (#)	A, B, C, E, J	12
***NOTHING**	C, D, G, I, J	11
***NEVER**	E, F, G	5
* **NO°**	C, G, J	5
NONE	A, F, G	4
NOTHING (fist to open at chin)	B, D	3
IMPOSSIBLE	E	2
BLOW-HAND	C	1
		Total token count: 125

(1) **From Video D**:
ASL: PAST HAVE PARENT (5 handshape) fs. SEE INFLUENCE $\overline{PARENT\ KICK-OUT}^{\text{headshake}}$ ME UNDERSTAND
English translation: For example, parent (with a P handshape) in the past it was signed like this (produces parent with a 5 handshape). I understand SEE influenced it and the sign became parent with a P handshape which signers rejected (negative headshake) and now we use the original form with the 5 handshape again.

(2) **From Video G**:
ASL: AFTER THAT WHATEVER PROCEED HIRE+++++ IX_{those} HEARING PROFESSIONAL WITH LITTLE-BIT MINIMUM SKILL SIGN AND UNDERSTAND DEAF CULTURE UNDERSTAND DEAF INNATE SO-FORTH **NEGATIVE HEADSHAKE**
English translation: After that, it seemed we hired anyone including hearing professionals who had little or minimal signing skills. Did they understand Deaf culture and many other aspects of our community? No (expressed with a negative headshake in its own timeslot).

(3) **From Video F**:
ASL: IX_{he} CAN HEAR ME (with a slight headshake) OBVIOUS ME **CAN'T** (no headshake) HEAR
English translation: He can hear while I obviously can't hear.

Genitive Adjacency in the Russian Noun Phrase as a Linearization Effect

Philip Shushurin, NYU

Introduction This paper suggests an explanation of adjacency effects between the head noun and core genitive arguments in Russian Noun Phrase (RNP) – henceforth **N GEN Constraint**.

(1) (*angličan) torgovlja **(angličan)** opiumom (*angličan)
(English.people.GEN) trade (English.people.GEN) opium.INS (English.people.GEN)
'the opium trade run by the English' [from Lyutikova (2016)]

The same constraint is seen with others such as Russ. *response parliament.*GEN *government.*DAT, *sketch artist.*GEN *to painting.*DAT etc. I propose that the constraint in (1) can be explained given (a) the requirement of such arguments to merge in the complement position and (b) inavailability of scrambling in such constructions (Kazenin 2007). I suggest that (a) can be derived from independently proposed principles of Harmonicity of Head Directionality (Dryer 1992) and the LCA (Kayne 1994), on the assumptions that in Russian, (c) the LCA doesn't target non-core arguments (see (Pylkkänen 2008) for core-noncore distinction) as well as PPs and adjuncts and (d) Harmonicity doesn't target phrases in agreement/concord relations with the rest of the structure. The proposed account can be generalized more generally across similar configurations beyond Russ., e.g. Case Adjacency (Stowell 1981, Neeleman 1994). **The word order in RNP** is subject to the following generalizations: (i) concording phrases must precede the noun (2a); (ii) the head noun must precede all non-concording material (2b); (iii) genitive core arguments must linearly follow the head noun (1).

(2) a. [Dimin-**a**] zarjadka (*Dimin-a)
Dima-ADJ.F charger.F (Dima-ADJ.F)
dlja ajfona (*Dimin-a)
for iPhone Dima-ADJ.F
'this Dima's iPhone charger' (=2b)

b. zarjadka [dlja ajfona] (*zarjadka)
charger for iPhone (charger)
[Dimy] (*zarjadka)
Dima.GEN (charger)

In contrast, non-core arguments, such as genitive possessors and oblique arguments occupy the position at the right periphery and may be linearly non-contiguous with the nominal head.

(3) a. kollekcija monet **professora**
collection coins.GEN professor.GEN
'professor's coin collection'

b. pis'mo Niny **bratu**
letter Nina.GEN brother.DAT
'Nina's letter to her brother'

Lyutikova (2016) proposes the following structure of (1), where the genitive argument is generated above the instrumental phrase and the head noun, and the head moves to the position above the genitive argument. However, she doesn't provide a clear motivation for head movement.

(4) torgovlja$_i$ [angličan [t$_i$ opiumom]$_{n'}$]$_{nP}$
trade English.people.GEN t opium.INS
'the opium trade run by the English'

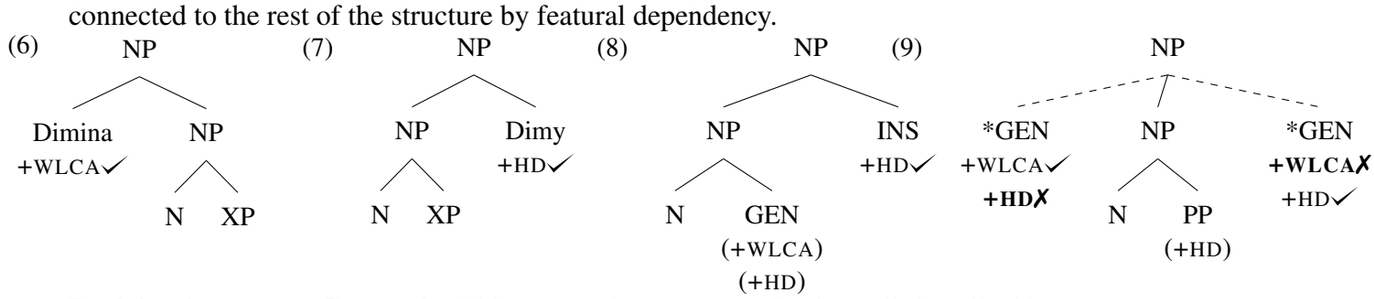
The gapping facts, however, indicate that the instrumental phrase is structurally higher (N+INS cannot be deleted, suggesting they do not form a constituent), contra Lyutikova's proposal. (Following Yoshida et al. 2012, I analyze Nominal Gapping as Ellipsis).

(5) a. torgovlja angličan opiumom po morju i ~~torgovlja angličan~~ čajem po zemle
trade English.p.GEN opium.INS by sea and ~~trade English.p.GEN~~ tea.INS by land
'the opium trade run by the English people by sea and the tea trade run by the English by land'

b. *torgovlja angličan opiumom po morju i ~~torgovlja~~ gollandcev ~~opiumom~~ po zemle
trade Engl.p.GEN opium.INS by sea and trade Dutch.p.GEN opium.INS by land
int. 'the opium trade run by the English by sea and the opium trade run by the Dutch by land'

Based on this evidence, I suggest that the genitive argument is in the complement position and the prepositional phrase merges to the right of the resulting NP. Since scrambling is impossible in such contexts, reordering of GEN and PP is impossible yielding N GEN Constraint. As for non-core arguments in (3), I suggest that they are right-merged to NP. This data raises the following questions. Why must modifiers precede the head? Why must non-concording dependents left-, rather than right-merge to the head? How can the N GEN Constraint be accounted for? **Proposal** *Assumptions* A1. I assume that LCA applies for some but not all types of phrases; more specifically, I suggest that LCA applies for core arguments and modifiers and doesn't apply for non-core arguments and other types of phrases (traditional adjuncts, appositives etc.). A similar position is found in Abels and Neeleman (2006). A2. Adopting ideas from Philip (2013), I suggest that Harmonicity of Directionality is a factor constraining linearization which applies to certain types of phrases, namely to phrases in featural dependency with the rest of the structure; A3. Following Norris (2014), I suggest that nominal heads do not agree with their arguments; what follows is that concord is the only type of featural dependency in the RNP. A4. HD applies for all types of Merge: left specifiers in a head initial languages count as a HD Violation, unless the merged phrase is featurally dependent. (5) Finally, following Norris (2014), I suggest that concord is upward valuation. I adopt a **weakened version of the LCA** (WLCA) as follows: If a phrase X is a core argument or featurally

dependent on the rest of the structure, then X must linearly precede all the nodes that it asymmetrically c-commands. Similar ideas are expressed in Brody (2000) and Brody and Szabolcsi (2003). **Harmonicity of Directionality** (HD) is defined as follows: All pairs of sister nodes X and Y, such that neither X nor Y is featurally dependent on the rest of the structure, must have the same directionality. Contra Costa (1996), Philip (2013), Lopez (2009) I suggest that Harmonicity and the LCA are not constraints in OT sense and do not participate in constraint interaction: violation of either rule results in ungrammaticality. However, different types of phrases are not subject to either WLCA or to HD. I assume linearization constraints to be surface filters, which apply after the syntactic derivation is complete. **Examples**. In the following examples, I will notate phrases which are subject to WLCA as +WLCA and phrases that are subject to HD as +HD 1. Modifiers. The modifier must be merged to the left, when merged with a nominal structure, t. Merging the modifier to the right would cause an LCA violation. Left-merge doesn't create a HD violation since the modifier is featurally dependent on the rest of the structure. (2a and 6). 2. Non-core arguments When a non-core argument is merged with a nominal structure, it must merge to the right. Merging the non-core argument to the left will cause a Harmonicity Violation. (2b and 7) 3. Core arguments When a core argument merges with a nominal head, it is linearized to the right, according to general Head-Initial Directionality parameter of Russian. This merge doesn't create a WLCA Violation, since the core argument merges with a head rather than a branching structure. Next, a non-core argument can be merged and linearized to the right, as a non-core phrase, this argument is not subject to WLCA. (1 and 8) 4. Non-concording core arguments can't merge to branching structures If a non-core argument is merged to a branching structure, (for example, a nominal head with a non-core argument in its complement position), either direction of merge will result in Linearization Violation: Left Merge would violate HD: unlike modifiers, the structure is not in featural dependency and cannot break Harmonicity; Right Merge would violate WLCA: as a core argument, it must right-merge with branching structures. (1 and 9) More generally, **core arguments which don't enter in agree/concord relations are predicted to be only permitted in the complement position**. Non-core modifiers such as intersective adjectives are subject to WLCA since they are connected to the rest of the structure by featural dependency.



Deriving *GEN GEN Constraint This proposal can account for the well-described ban on two genitive arguments with core argument semantics, cf. Russ. *performance aria.*GEN **Chaliapin.*GEN/ ✓*Chaliapin.*INS 'performance of the aria by Ch.' (Pereltsvaig (2017) a.o.). Since both arguments are core, they are both subject to WLCA. The higher of those arguments cannot be merged for the same reasons as the genitive argument in (9): each direction of merge will create a violation of either WLCA or HD. By assumption, instrumental phrases are not core arguments and thus are not subject to WLCA: they can be merged to the right. **N GEN Requirement and Case Adjacency** The analysis proposed here can be further extended to the observation by Belk and Neeleman (2017) that certain case-marked objects must be linearly adjacent to verbal heads (Case Adjacency) in many head-initial languages. Such effects are not found in Head-Final Languages. Assuming that Case-Adjacent arguments are core and non-agreeing, they are predicted to be disallowed in non-complement positions for the same reasons as Genitive Arguments in (9). BN's account cannot be directly applied to DP-internal genitive arguments given the widespread view (e.g. Pesetsky (2013)), that Genitive Case in the RNP is not assigned by the nominal head but has a different source. **The Absence of Case Adjacency in Head-Final Languages** can be explained if one adopts the view (Ernst 2003) that Head Directionality is parameterized for head-complements (i.e. dependents merging with a head) but not traditional specifiers, that is dependents left-merging with branching structures. If a language allows head-final order at the first merge, it means that left-merge of a dependent would never violate HD, since such structure would be harmonic. Phrases that are subject to both WLCA and HD in HF languages would be able to merge with heads as well as branching structures. Another possibility is that merging Direct Objects is preferred in the complement position by independent (i.e. semantic) reasons, but scrambling can reorder the surface order in HF languages but not in (many) HI languages. The assumption then is that HD and the WLCA apply as filters after the scrambling has taken place. **Implications for FOFC Effects** Adopting the system proposed here and assuming that types of phrases that are exempt from HD

and WLCA in different languages is subject to parametric variation, FOFC violations can be seen as violations of both HD and WLCA.

FREQUENCY-DRIVEN ASYMMETRIES IN MORPHOLOGICAL DECOMPOSITION AS REVEALED BY MASKED VISUAL PRIMING

Roberto Petrosino & Jon Sprouse – *University of Connecticut*

INTRODUCTION. A large body of masked priming studies has shown that, before being visually recognized, words are decomposed in smaller units that seem to correspond to morphemes [1, 2]. This procedure, called *morphological decomposition*, (a) occurs in words that are, or seem to be, made of more than one morpheme (morphologically transparent words: e.g., *driver→{drive}-{er}*; morphologically opaque words: e.g., *archer→{arch}-{er}*, even though *archer* is not someone who arches); but (b) it does not occur in words that contain a root plus a non-affixal ending (e.g., *brothel↛{broth}-{el}*, where *el* is not an English suffix). Two sets of models have been proposed to explain these facts. On the one hand, *single-route, decomposition-first models* (1-DMs) argue that a word is recognized through an obligatory procedure of decomposition *automatically* occurring before lexical access [3]. On the other hand, *dual-route race models* (2-RMs) argue that a word is recognized through the fastest of two parallel processing routes racing against one another: the *parsing route*, in which the word is *automatically* decomposed into its constituents; and the *storage route*, in which the word is searched as a whole (i.e., without being decomposed) via lexical access [4]. Setting aside the evident differences, both models assume decomposition to be a stripping mechanism that is *automatic*, namely that occurs independently of lexical information such as frequency.

(1) FREQUENCY-FREE DECOMPOSITION HYPOTHESIS
Morphological decomposition is *not* affected by word frequency.

We test the prediction (1) by eliciting masked priming in English regular plural forms like *pixels* and *worlds*. Lexical decision studies have shown that, in many languages (among which Dutch and English), PLURAL-DOMINANT plural forms such as *pixels*, whose surface frequency is higher than the surface frequency of the singular form *pixel*, are visually recognized faster than SINGULAR-DOMINANT plural forms such as *worlds*, whose surface frequency is lower than the surface frequency of the singular form *world* [4]. In compliance with (1), both 1-DMs and 2-RMs predict that both plural forms trigger priming, though for different reasons. In 1-DMs, since the decomposition procedure is obligatory, both forms are expected to decompose and trigger priming. In 2-RMs, since both routes are always triggered, both forms are expected to decompose in the parsing route and therefore trigger priming.

MATERIALS & PROCEDURE. Four conditions were tested (20 pairs each). (1) The *sgdom-sgdom condition* consisted of singular-dominant primes and singular-dominant targets (e.g., *worlds-HEAVENS*). (2) The *pldom-sgdom condition* consisted of plural-dominant primes and singular-dominant targets (e.g., *pixels-GODS*). We made sure that all plural forms in conditions (1-2) all had the [z] allomorph of the English plural suffix, so to avoid potential confounds due to the phono-orthographic differences with the other allomorphs (i.e., [IZ, S]). The remaining two conditions were added as sanity-check conditions: (3) the *transparent condition* consisted of pairs carrying a semantically and morphologically transparent relationship (e.g., *boneless-BONE*); and (4) the *orthographic condition* consisted of pairs in which the primes (e.g., *banjo*) were made of the corresponding target word (*BAN*) plus an additional non-affixal ending (*-jo*). One hundred and thirty-one participants were recruited through Amazon Mechanical Turk and took the experiment online through PsychoJS [5]. Subjects were asked to read the capitalized letter strings on the display and decide as quickly and accurately as possible whether or not each string is a word. The primes were preceded by a forward mask ('#####') and followed by the target word. The forward mask was 500ms-long, while the prime was 34ms-long.

RESULTS & STATISTICAL ANALYSES. Fig. 1 reports mean priming effects and Cohen's *d* (ES, effect size) for each condition over the bars. Response times for each condition were fitted into a linear mixed-effect regression (LMER) model and a Bayes Factor (BF) model. Each model had *log RT* as the dependent variable, RELATEDNESS (i.e., related/unrelated prime) as fixed factor, and SUBJECT and ITEM as random factors (intercepts only). Fig. 1 reports *p*- and $BF_{1,0}$-values for each condition under the bars. The LMER analysis showed significant priming effects for both the sgdom-sgdom condition and the transparent condition (ps$< .001$); and non-significant effects for the pldom-sgdom condition and the orthographic condition (ps$\geq .20$). For the interpretation of the BFs, we refer to Jeffreys (1961)'s interpretive table. The BF analysis confirmed that the sgdom-sgdom and the transparent conditions substantially supported the alternative hypothesis ($BF_{1,0} > 3$); and the pldom-sgdom and the orthographic conditions strongly supported the null hypothesis ($BF_{1,0} < 0.1$).

We then Dunn-corrected pairwise comparisons, and the following combinations were found significant: sgdom-sgdom vs. pldom-sgdom, transparent vs. pldom-sgdom, transparent vs. orthographic (ps= .003); the comparison sgdom-sgdom vs transparent was not significant ($p = 1$). Uncorrected BF comparisons confirmed these results.

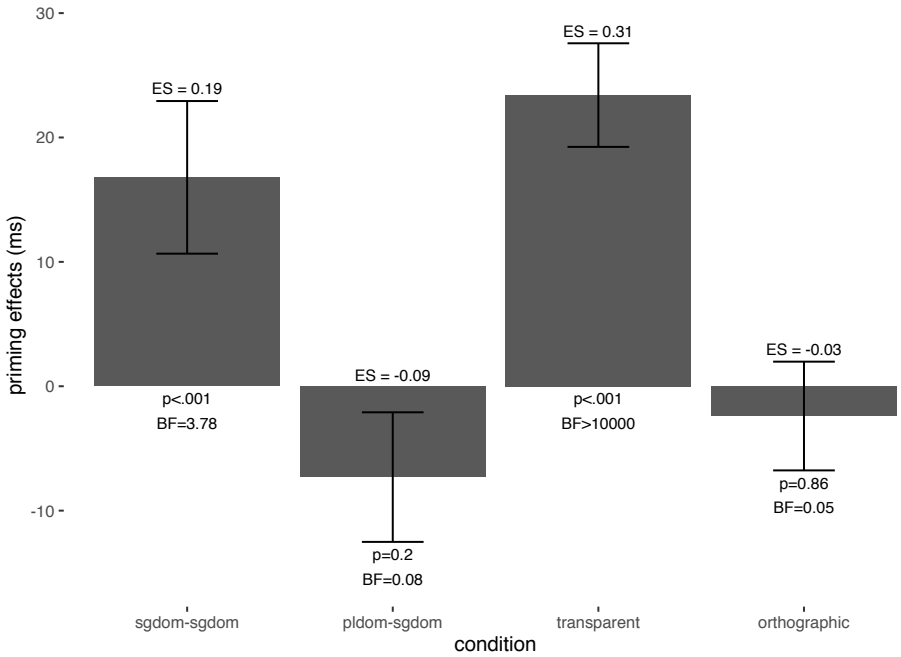


Figure 1

DISCUSSION. Our results showed that singular-dominant plural forms like *worlds* prime, whereas plural-dominant plural forms like *pixels* do not. These results challenge both 1-DMs and 2-RMs. Both models wrongly predict that PLURAL-DOMINANT and SINGULAR-DOMINANT plurals prime and decompose, since they assume a procedure of decomposition that is not affected by frequency (1). These results could be explained if, *contra* (1), we assume decomposition may be sensitive to the frequency-based asymmetries between the two alternants of the number paradigm. When a singular-dominant plural form (*worlds*) is presented, it is decomposed as world-s and the singular form (*world*) is recognized right away because it is high-frequency, and therefore triggers priming facilitation effects. When a plural-dominant plural form (*pixels*) is presented, it starts being decomposed but the singular form (*pixel*) takes time to get recognized because it is low-frequency; as a consequence, it does not trigger priming facilitation effects onto the target.

CONCLUSIONS. In this paper, we presented novel experimental data that tested the sensitivity of decomposition to frequency-based asymmetries between singular and plural forms in English. To this end, we elicited suffix priming in two kinds of plurals: PLURAL-DOMINANT plurals (*pixels*), which are more frequent than the corresponding singular form (*pixel*); and SINGULAR-DOMINANT plurals (*worlds*), which are less frequent than the corresponding singular form (*world*). In our experiment, significant priming effects were elicited by singular-dominant plurals, but not by plural-dominant plurals. This asymmetry was not expected in either of the two models considered (i.e., the 1-DMs and the 2-RMs), since both assume that decomposition is not affected by frequency according to (1). While not specifically supporting either model, the results reported above seem to suggest that decomposition of English plural forms is affected by frequency: in singular-dominant plurals, decomposition is faster because the singular form is high-frequency and is therefore readily identified; in plural-dominant plurals, decomposition is slower because the singular form is low-frequency and therefore takes more time to be identified. Further research is expected to be undertaken to clarify the mechanisms and the conditions under which frequency affects decomposition at early stages of processing.

REFERENCES. [1] K. Rastle et al. *Lang. Cogn. Process* 15 (2000). [2] K. Rastle et al. *Psychon. Bull. Rev.* 11 (2004). [3] M. Taft and K. I. Forster. *J. Verbal Learning Verbal Behav.* 14 (1975). [4] R. H. Baayen et al. *J. Mem. Lang.* 37 (1997). [5] J. Peirce et al. *Behav. Res. Methods* 51 (2019).

Factors involved in limited convergence to an interviewer in (ING) variation

As a well-documented, stable, internally and externally conditioned variable, variable (ING) (e.g. *workin' ~ working*; Tamminga 2014) provides fertile ground for exploring the intersections of social and cognitive factors in speech. In this study, we use (ING) to probe cross-speaker convergence effects among speakers of Philadelphia English. While our results are overwhelmingly null, our study still highlights multiple avenues for future research.

Background. Studies of interspeaker alignment in conversation tend to draw from two fundamental perspectives: Communication Accommodation Theory, or CAT (Giles & Ogay 2007); and theories of repetitive speech perception and production (Pardo 2013; Tamminga 2014, 2019). CAT applies ideas of identity-building to explain why speakers often end up speaking more or less similarly to each other over the course of an interaction (Gallois et al. 2005). In contrast, speech perception/production theories argue that speakers tend to repeat recently heard variants during a conversation (summarized Casserly & Pisoni 2010). Recent experimental accommodation findings suggest that cognitive aspects of speech perception and production are indelibly connected to social factors (Pardo 2006, 2013; Aguilar et al. 2016; Walker & Campbell-Kibler 2015). By using corpus data, this study draws on natural speech to attempt to expand our understanding of socially mediated cognition.

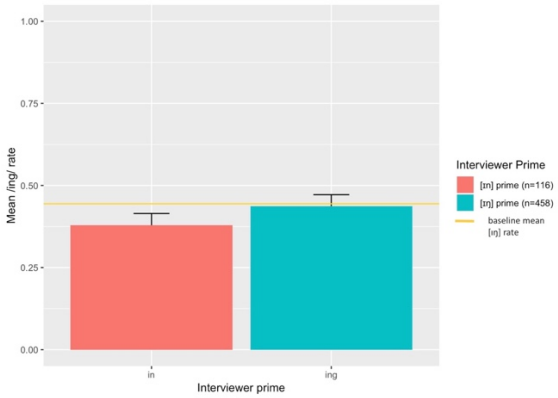
Data. Data come from 122 speakers of the Philadelphia Neighborhood Corpus (Labov & Rosenfelder 2011). Every instance of speaker (ING), and certain instances of interviewer (ING), were coded for pronunciation (apical [ɪn] or velar [ɪŋ]). Additional internal and external predictors (e.g. token grammatical category, speaker gender) were also coded. Following Tamminga (2014), instances of interviewer (ING) were coded when they immediately preceded segments of speaker (interviewee) speech.

Analysis. Data were analyzed at the aggregate and individual speaker levels to see if speakers produced more [ɪŋ] following instances of interviewer [ɪŋ], relative to their baseline levels of [ɪŋ]; and if speakers produced more [ɪn] following instances of interviewer [ɪn], relative to their baseline levels of [ɪn] (Figures 1, 2). Further analyses examined the effect of speaker gender and token grammatical category on the presence of convergence/priming patterns (Figures 4, 5). Statistical testing is completed with chi-squared tests and a linear regression (Table 1).

Results. Our first analysis did not find statistically significant differences between aggregate speaker [ɪn]-primed mean [ɪŋ] rates and aggregate speaker baseline mean [ɪŋ] rates, nor between aggregate speaker [ɪŋ]-primed mean [ɪŋ] rates and aggregate speaker baseline mean [ɪŋ] rates. Our second analysis confirmed this null effect of interviewer [ɪn] and [ɪŋ] primes on individual speaker mean [ɪŋ] rates relative to individual speakers' baseline mean [ɪŋ] rates. Thus, we cannot make any claims about cross-speaker convergence/priming effects in our data. Surprisingly, it appears that cross-speaker accommodation patterns are different from intraspeaker persistence patterns (Figure 3), but the source of these differences is unclear. We find expected significant gender effects, such that female [ɪŋ] rates are significantly higher than the mean of male and female [ɪŋ] rates. However, these gender effects do not reveal anything remarkable, as we still cannot say how the gender effect might interact with accommodation trends. In terms of grammatical effects, we see tentative evidence that monomorphemic and progressive (ING) are not behaving the same way, yet small token numbers limit the claims we can make on this front.

Implications. Our overwhelmingly null results leave open many avenues for future research. Overall, corpus studies do appear to be valid and necessary for investigating accommodation effects. Furthermore, variable (ING) still appears relevant for expanding our understanding of linguistic phenomena. Any extensions of this study would benefit from more data.

Figure 1. Aggregate speaker mean [ɪŋ] rates || **Figure 4.** Aggregate speaker mean [ɪŋ] by gender



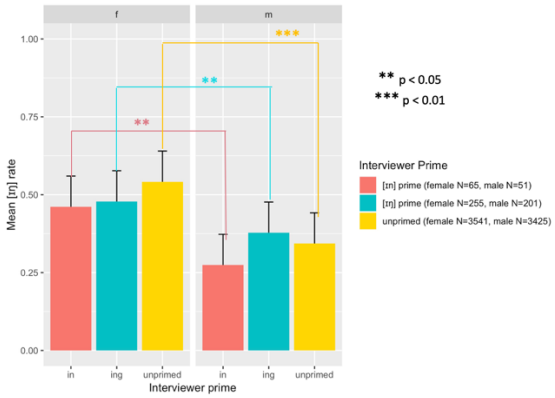
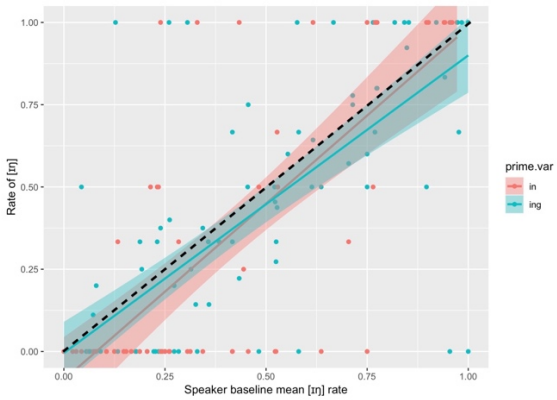


Figure 2. By-speaker values for (ING) || **Figure 5.** Effect of progressive primes on prog/mono targets



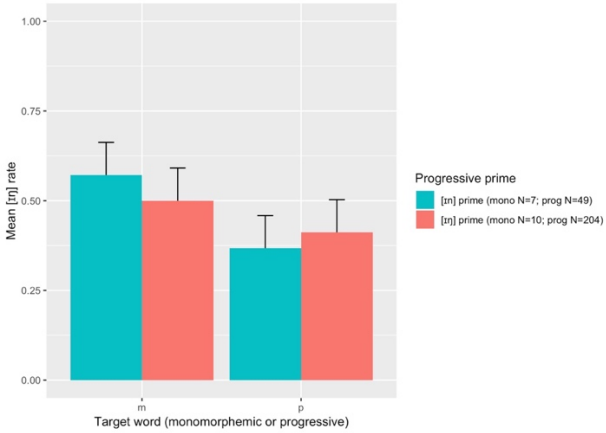
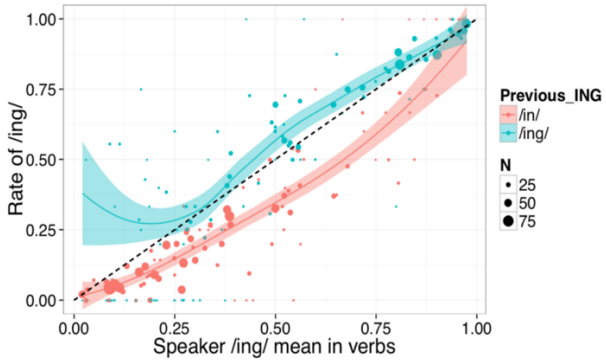


Figure 3. Intraspeaker persistence (Tamminga 2014:103) ||



Selected References.

Aguilar, L. et al. (2016). A Dyadic Perspective on Speech Accommodation and Social Connection • **Campbell-Kibler, K**. (2010). New Directions in Sociolinguistic Cognition • **Casserly, E. & Pisoni, D**. (2010). Speech perception and production. **Gallois, C., Ogay, T., & Giles, H**. (2005). Communication Accommodation Theory: a look back and a look ahead • **Giles, H., & Ogay, T. (2007)**. Communication Accommodation Theory • **Labov, W. & I. Rosenfelder**. (2011). The Philadelphia Neighborhood Corpus. • **Pardo, J.** (2006). On phonetic convergence during conversational interaction • **Pardo, J. S.** (2013). Measuring phonetic convergence in speech production • **Tamminga, M.** (2014). *Persistence in the Production of Linguistic Variation* • **Tamminga, M.** (2019). Sources of Microtemporal Clustering in Sociolinguistic Sequences • **Walker, A. & Campbell-Kibler, K.** (2015). Repeat what after whom?

Table 1. Linear Regression Output *(categorical predictors sum-coded, estimates in log-odds)*[1]

	Estimate	Std. Error	z value	p-value
(Intercept)	-0.7	0.2	-2.9	0.00 **
Prime variant [ɪn]	0.0	0.2	0.0	0.97
Gender (female)	0.5	0.2	2.1	0.04 *
Prime variant [ɪn] * gender (female)	0.1	0.2	0.5	0.59

[1]: In table, * = $p < 0.05$ and ** = $p < 0.01$