JS Journal of Historical Syntax
Volume 10, Article 8: 1–46, 2026 · DOI: 10.18148/hs/2026.v10i8.286

MODELING THE ACQUISITION OF LABILE VERBS IN MIDDLE ENGLISH*

Jordan Kodner
Stony Brook University

Achim Stein
University of Stuttgart

Carola Trips
University of Mannheim

Abstract This paper applies current mathematical models of learnability (Yang 2005, 2016, Irani 2019, Li 2024) to historical data to gain a precise understanding of how children in the past learned formal aspects of language. We take the learnability of the causative-anticausative alternation as an example (Irani 2019) and address the question of how children learned the class of labile verbs in Middle English. We investigate the contact hypothesis, which argues that the well-known influx of Old French verbs of change-of-state coming into Middle English (Ingham 2020) had an effect on learning the alternation as a productive rule. Drawing on work on learnability in diachrony (Kodner 2019a, 2020, 2023b, Trips & Rainsford 2022), we present a study which applies the Tolerance and Sufficiency Principle(s) (TSP; Yang 2016) to the lexicons of Present-Day English first-language learners and to the lexicon of Middle English change-of-state and motion verbs. We evaluate the conditions under which learners would posit that all of these verbs are potentially labile using frequency-weighted sampling. Significantly, we show that both morphological change (loss of distinct causative morphol-

* We would like to thank the audiences at the 57th Annual Meeting of the Societas Linguistica Europaea (SLE) in Helsinki, 21-24 August 2024, and at DiGS 26 in Oxford, 23-26 June 2025. Thomas Rainsford (Stuttgart) contributed the first simulations of "Middle English learners" to the talk held at DiGS in Oxford. Cristina Gutiérrez Alcalá (Stuttgart) provided the decision-tree based classification of intransitive constructions for the exploration of Modern English CHILDES data in Section 4.3. Charles Yang provided valuable feedback on how (not) to apply the Tolerance and Sufficiency Principle(s). Work presented here was funded by the Deutsche Forschungsgemeinschaft (DFG) for project H2 *Change and acquisition of verbal structures*, and project H3 *Changing the structure of the input in contact: verbal prefixes and verb particle combinations* as parts of the Research Unit FOR 5157 "Structuring the input in language processing, acquisition and change" (SILPAC), grant no. 437487447.

©2026 Kodner & Stein & Trips
This is an open-access article distributed under the terms of a Creative Commons License (creativecommons.org/licenses/by/4.0).

ogy) and contact (copying of French verbs) contribute to the rise in lability. Overall, we present a new methodological pathway for historical linguistics: the combination of historical corpus linguistics with cognitively motivated computational modeling to reveal how the historical child learned grammatical rules.

1 Introduction

The *causative alternation* (or causative–inchoative or causative-anticausative alternation) is a lexical alternation whereby a single verbal lexeme participates in two argument-structure frames: a transitive, causative frame (*X causes Y to change*) and an intransitive, inchoative one (*Y changes*) (compare Present-Day English (PDE) *I broke the window*/*The window broke*). The phenomenon is defined among others by Levin (1993), Levin & Hovav (1995), Schäfer (2008). Languages differ in how the two variants are morphologically related (causativizing, anticausativizing, or non-directed).

From a diachronic perspective, this alternation is an intriguing phenomenon: compared to the other Germanic languages, Present-Day English exhibits the highest proportion of verbs which undergo this alternation without changing the verb form, a property that has been called 'lability' by Haspelmath (1993) (see also Comrie 2006). Although Old English (OE) showed some labile verbs, the causative-anticausative alternation was mainly expressed morphologically by affixation (e.g. *drygan-**a**drygan* 'dry,' and/or by stem-vowel change (e.g. *sencan-sincan* 'sink,' cf. García García 2019). However, these distinctions were mostly leveled out in ME, resulting in more verbs with zero morphological alternation (i.e., labile patterning, see e.g. van Gelderen 2018). Apart from this explanation, which sees internal factors as the primary driving force of this change, the language contact hypothesis has also been proposed to explain the increase in lability in English. Specifically, authors like Ingham (2020) have argued for language contact with Old French (OF) (see also García García & Ingham 2023). Ingham notes that OF Change-of-State (CoS) verbs were labile at the period of contact via intense bilingualism, and that by the end of the 15th century, this class of verbs in ME had reached 75% of lability. Ingham implicitly addresses language acquisition here, adopting the argument that the enrichment of the lexicon with OF labile verbs and their propagation in the speech community must have had acquisition as its source.

It has long been argued that children's novel productions during language acquisition play a causal role in language change (Paul 1880, Kiparsky 1965, Lightfoot 1979, Niyogi & Berwick 1997, Yang 2002, Kroch 2005, Labov 2007,

van Gelderen 2011, Cournane 2017, Kodner 2020, *i.a.*).[1] Under this thesis, the changes that occur in a language reflect some of the innovations produced by child learners in the past. We should then expect that the observed overgeneralizations among learners of PDE regarding the causative alternation (Bowerman 1974, 1982, Bowerman & Croft 2008) may resemble the sort of over-extensions that led to the development of the alternation in the history of the language. As the Tolerance and Sufficiency Principle(s) (TSP; Yang 2005, 2016), a quantitative model of productivity learning, has a successful track record in providing mechanisms for such changes (Björnsdóttir 2021, Trips & Rainsford 2022, Kodner 2023b, Yang 2024, Dolatian & Kodner in press), we will leverage them here, from a historical perspective to ME data, as a step towards uncovering the mechanism by which the ME child learned the event structure of labile verbs and lability achieved productivity over time in the speech community. By adopting a quantitative approach and a cognitively motivated notion of productivity informed by research on child language acquisition, we will test Ingham's hypothesis regarding the influence of French-based verbs more directly as well as a potential causal role of morphological leveling increasing the number of apparently labile verbs on the rise of productive lability.

The outline of the article is as follows: Section 2 establishes the theoretical definitions of the causative–anticausative alternation and outlines the event-structural parameters that constrain lability. Section 3 provides the historical and contact-linguistic background of the causative–anticausative alternation and presents a corpus study building on our previous etymological annotations of ME corpora to contrast native and French-copied verbs. Section 4 establishes the learnability framework; it provides an empirical baseline by applying the TSP to first-language acquisition data (CHILDES), where we expand our focus from CoS verbs to include the second major labile class, motion verbs. Section 5 returns to the historical contact situation, presenting quantitative simulations that test the interaction between morphological leveling and the French lexical influx in Middle English. Finally, Section 6 discusses the implications of these findings for theories of language contact and acquisition-driven change.

1 Novel childhood productions are often described as "errors." This is a fraught concept, particularly in the context of language variation and change, so we adopt the terms "novel productions" and "innovations" instead.

2 Theoretical foundations of lability and event structure

The term lability is generally used to denote a verbal property, i.e., for lexemes which participate in an alternation without any formal change. In typological studies (e.g., Haspelmath 1993), a verb is considered to show Patient lability if it is attested in the causative-anticausative alternation, as opposed to Agent lability, where the object is omitted and the Agent remains in subject position (cf. Letuchiy 2010 for an overview; see also Section 4.3). Regarding the form of the verb, Letuchiy (2010: 238) makes reference to Nedjalkov's (1969) typological distinction in morphological marking, of which "labile" is the one "without any morphological change." We therefore define the causative-anticausative alternation as the wider notion, including cases that involve overt morphological markers such as prefix or vowel change, and lability as the narrower notion, where "isomorphous" verb forms (we prefer to use "same-form") are used in both frames (e.g. *Tom broke the window* / *The window broke*).

More often than in the Germanic languages, Romance languages mark the anticausative variant of the verb with reflexive clitics such as *si* in Italian or *se* in Modern French, as in *La vitre se casse* 'The window breaks.' For contact with Old French (OF), this pattern is less relevant, since the reflexive marker, despite being attested, develops only later into a systematic marker of unaccusatives (see Heidinger 2010 and Inglese 2025 for diachronic studies of French and Italian, and Percillier, Schauwecker, Stein & Trips 2024: ch. 2.5 for reflexives in ME under French contact influence, see also section 3.1). OF thus shows a higher degree of lability in the narrow sense than Modern French.

In languages with historical causative or anticausative markers, including members of Germanic and Romance, the emergence of Patient lability is often the synchronic result of morphological leveling. We use the term "leveling" here (and in particular in Section 5) in the sense of analogical change where morphological distinctions between related stems or derivations are reduced or eliminated. In the context of the causative-anticausative alternation, leveling results in same-form verbs, i.e., lemmas where the causative and anticausative uses are phonologically identical. It is these leveled, identical forms that provide the primary positive evidence for a productive labile rule. We will use both terms, "lability" as a grammatical property of verbs or, more generally, of the target grammar (the rule that the learner acquires), and "leveling" to denote the formal change.

The event structure of alternating verbs consists of a causing and a resulting event (Pinker 2013 [1st ed. 1989], Levin 1993). The anticausative construction drops the causing event. The causative-anticausative alternation

is blocked if semantically relevant components like Manner or Control are linked to the causing component (CAUSE) of the causative construction. According to Pinker (2013: 271), "[v]erbs of force exertion such as *push* or *shove*, [...] always specify a manner within the ACT structure", i.e. within CAUSE. As a consequence, the unaccusative construction is grammatical with *move* and *pour*, but not *push*, *hit*, *cut* and other verbs expressing manner of the action:

(1) a. *move*: [x CAUSE [y MOVE]]
b. *pour*: [x CAUSE [y MOVE [by POURING]]]
c. *push*: [x CAUSE [y MOVE] [by PUSHING]]

The fact that Manner blocks the anticausative construction is expressed in Haspelmath's condition "absence of agent-oriented meaning components" (Haspelmath 1993: 92-93, without reference to Pinker). From a typological point of view, Haspelmath (1993: 100-103) considers English as quite unique in showing a strong preference for lability, also compared to the other Germanic languages. For example, German shows morphological marking, i.e. transitivizing prefixes as in *versenken* (2), or the reflexive marker *sich* in the intransitive anticausative use of *öffnen* (3), compared to their English labile equivalents *sink* and *open*.

(2) *Sie versenkte das Boot — Das Boot (ver)sank*
she sank the boat the boat sank
'She sank the boat. — The boat sank.'

(3) *Sie öffnete die Tür — Die Tür öffnete sich*
she opened the door the door opened
'She opened the door. — The door opened.'

Lability in PDE is a well studied phenomenon, but its definition is subject to different weighing of syntactic, semantic and morphological criteria. McMillion (2006) is the most comprehensive work on the topic noting that there are over 800 labile verbs in PDE, thus corroborating Haspelmath's observation that English is exceptional in this respect in the set of the Germanic languages. For McMillion (2006) the "switch in causativity" (p. 1), i.e. the observation that "... the transitive and intransitive patterns differ only in the dimension of causativity" (p. 21) is the central criterion for distinguishing labile from other similarly alternating verbs. For us, this is simply the causative alternation (e.g. *kill* vs *die*) where the criterion of "same surface verb form" is not relevant. At some later point McMillion, however, also states that the term labile refers to verbs that undergo the causative alternation where the transi-

tive causative and intransitive inchoative argument structures use the same verb form (p. 9). For us, this is the narrow definition of the phenomenon. Concerning the lexical semantics of these verbs McMillion, includes inherently process-oriented verbs (*flash*, *bounce*), as well as verbs that encode an inherent result (*break*, *melt*) which make up the majority of labile verbs in his study. The result may be CoS verbs, change-of-location verbs, verbs denoting aspect, or other properties (p. 204). For us, the presence of a causal Agent in the transitive argument structure and its absence in the intransitive argument structure is crucial. This excludes many of the verbs in McMillion's list. Based on McMillion's wider definition, it is not surprising that he finds this large amount of labile verbs (which comprise core labile verbs (*melt*, *sink*), and peripheral cases such as *gallop the pony*, p. 21). For our studies of child-data and historical data we will use the term lability only for verbs with the same surface form in transitive causative and intransitive anticausative argument structures where there is a clear causal Agent in the former and its clear absence in the latter. We will focus on CoS verbs in the historical study (Section 3) and add the class of motion verbs in the acquisitional study (Section 4.3).

3 Lability in the History of English

As mentioned in the introduction, Old English, like the other Germanic languages, only had a small set of labile verbs: Visser (1963: 99) counts fifty-five verbs of the "double-function or amphibious" type, among them *brecan* 'break,' *gladian* 'gladden' and *openian* 'open.' McMillion lists 110 verbs (for a discussion see McMillion 2006: ch 7.3). There is agreement that the much more common way to mark the causative alternation was by morphological differentiation. The most important morphological means for expressing the alternation was affixation, more precisely, by reflexives of the Proto-Germanic *-*j*(*a*) suffix that was often added to strong, intransitive (unaccusative) verbs (e.g. van Gelderen 2018, García García 2019). In OE, we find pairs of strong base verbs and weak derived verbs, some examples are given in (4) (cf. García García 2019)[2]. These forms often exhibit opaque stem vowel alternations which have their origin in regressive assimilation from the *-*j*(*a*) suffix (only (4-e) shows an overt -*i* suffix). Processes of syncope led to the alternation becoming opaque (e.g. Ringe & Taylor: 264-273) :

(4) a. *sincan – sencan*

2 Compare Modern German: (a) *sinken* 'sink' – *versenken* 'cause to sink,' (b) *bitten* 'to beg' – *beten* 'pray,' (c) *trinken* 'drink' – *tränken* 'cause to drink,' (d) *fallen* 'fall' – *fällen* 'cause to fall,' no equivalent for (e).

b. *bītan – bǣtan*
c. *drincan – drencan*
d. *feallan – fyllan*
e. *faran – ferian*

Apart from stem-vowel changes there was also the possibility to use aspectual prefixes to transitivize and causativize verbs, predominantly *ā-* (see e.g. Dietz 2004, Elenbaas 2007). Examples include *drygan* 'dry' – *ā-drygan* 'make dry,' *belgan* 'swell with anger' – *ā-bylgan* 'anger, offend,' *steorfan* 'kill' – *ā-styrfan* 'cause to die.' We call the type that shows stem-vowel changes and affixation the Germanic type of marking the causative alternation.

In ME, stem vowel alternations and morphologically derived patterns were gradually leveled leading to a situation where some new labile forms arose, (5-a) and (5-b), though a small set of verbs of the Germanic type was retained (5-c). Further, some OE labile verbs were retained (5-d), some became labile (5-e). The language also gained labile verbs as copies from OF (5-f) (cf. Ingham 2020).

(5) a. *sinken* 'to become submerged, to cause to submerge'
b. *drien* 'to become dry, to make dry'
c. *quellen* 'kill' – *quelen* 'die' (> OE cwellan – cwelan)
d. *breken* 'to go to pieces, to cause to break'
e. *cleven* 'to become split, to split' (> OE (to)cleofan, trans.)
f. *avauncen* 'to assist, to be profitable'

So the main picture is a decreasing number of verbs of the Germanic type and an increasing number of labile verbs. Regarding the cause of this increase the literature discusses two options: internal and external factors. Proponents of the internal change hypothesis like McMillion (2006: 204) see "[t]he lack of middle markers and the explicitness of reflexive markers" as having contributed to the increase of labile verbs whereas van Gelderen (2018) assumes that the causativizing suffix, underlyingly *-i* (< Proto-Germanic *-*j*(*a*)), is no longer transparent by late OE (see also García García 2019) and verbs are reanalyzed as labile by the learner "...with lability for the unaccusative-causative alternation as well as lability for the intransitive-transitive one" (p. 111). Proponents of the language-contact hypothesis like Ingham (2020: 447) assume that verbs copied from OF "provided models for the generalization of lability across these English verb classes" (see also García García & Ingham (GGI) 2023). Since we are interested in how the historical child learned labile verbs in ME and whether contact with OF had an influence on it, we will take a closer look at this contact situation and Ingham's (2020) and GGI's (2023)

studies in the next section.

3.1 *The contact hypothesis as an explanation for the rise of labile verbs in Middle English*

There is a large body of work on the contact situation that followed the Norman Invasion in 1066 with a focus on the massive influx into the English lexicon (Rothwell 1980, Kibbee 1991, Short 1995, 2007, Ingham 2012, Durkin 2014) but also more recently on structural copying, especially the copying of verb argument structure (see the papers by Ingham and GGI cited above and Percillier et al. 2024). Since this contact situation prevailed for more than three centuries, it underwent different stages. The first was a period of diglossia that lasted for about two to three generations with (continental) OF as the language of the invading ruling class and ME as the native language of the main part of the population. From the end of the 12th century to the middle of the 14th century, a stage of bilingualism has been assumed where English and French co-existed in educated bilingual individuals, so the nature of contact shifted from societal to individual bilingualism (Ingham 2012, Ingham 2020). More precisely, the insular contact variety of Anglo-French[3] had emerged by that time and was used as a language of record-keeping and a medium of instruction in education. Probably due to the Black Death and the concomitant death of many teachers (clergymen), Anglo-French lost its status as a vehicle language in school and as a medium of oral and written communication, and at the end of the 14th century, the final period emerged, when it became a foreign language that was acquired through conscious learning. Ingham assumes that at this point the transmission of Anglo-French had collapsed and it was not very long before it disappeared from the scene in Britain.

We are particularly interested in the middle stage of bilingualism because it is well-established in the literature on language processing and acquisition that "contact in the individual" facilitates cross-linguistic influence (CLI) (e.g. Müller 2003, Bailey, Lockary & Higby 2024). We contend that this extends to the *historical* bilingual, encompassing the level of syntax, and, specifically, the mapping of argument structure.

Coming back to the phenomenon under scrutiny here, Ingham (2020) and GGI (2023) claim that the labile argument structure of verbs copied from OF was replicated in ME by the historical bilingual: when OF verbs were copied to ME their argument structure was copied along. Studies like Heidinger (2010) show that OF exhibited many labile verbs (CoS, motion) and thus support this assumption. Ingham (2020) gives a number of concrete ex-

3 We will use "OF" as a cover term for continental Old French and Anglo-French.

amples of this type of grammatical replication (cf. Heine & Kuteva 2005): In the first case, French-style argument structures were created for the ME verbs that replicated the structures of the OF verbs (see Table 1). For example, the argument linking of OF CoS and motion verbs *roster* 'roast' and *rouler* 'roll' is the same for the ME copies *rosten* and *rollen* (cp. Ingham 2020: 461). According to Ingham and GGI these cases show a wholesale transfer of the two argument structures of the OF labile verbs.

OF	*rostir* ('roast')	{NP	_____	NP }
	rouler ('roll')	Agent		Patient/Theme
		{NP	_____ }	
		Patient/Theme		
ME	*rosten* ('roast')	{NP	_____	NP }
	rollen ('roll')	Agent		Patient/Theme
		{NP	_____ }	
		Patient/Theme		

Table 1 Grammatical replication of OF labile verbs (Ingham 2020:461)

Some examples for the OF labile verb *roster* and its ME equivalent *rosten* are given here:

(6) OF

a. *il* ***rosterent*** *phase* (= *a sacrificial lamb*) *sur le feu*
they roasted lamb over the fire
(MED, Bible Royal 300vb)

b. *mettez le gars sur un espé* [...] *e lessez* ***roster*** *molt bien ...*
make the duck on a stick and let roast very good
(MED, Five Med MSS 117.E1)

(7) ME

a. *By a mykel fir he sat,* ***Rostyng*** *a swyn gret & fat.*
by a big fire he sat rosting a pig great and fat
(MED, a1450(a1338) Mannyng Chron.Pt.1 (Lamb 131)12342)

b. *Turbot ... when it* ***rosts****, springle on salt.*
Turbot ... when it roasts sprinkle on salt
(MED, ?a1475 Noble Bk.Cook.(Hlk 674)97)

Under the assumption that for bilinguals lexical access in one language may affect, or be affected by, activation of words in another language, such CLI effects facilitate lexical access (see, e.g., Bailey et al. 2024 and the sources cited therein). In the case at hand, lexical access to the argument structures of verbs from both contact languages is facilitated in the historical bilingual and verb-based syntactic transfer can take place from OF verbs to the corresponding

ME verbs (see Table 2):

OE	*awacian* ('awake')	{NP Patient/Theme	_____ }	
ME	*awaken* ('awake')	{NP Agent	_____	NP Patient/Theme }
		{NP Patient/Theme	_____ }	
OF	*esveiller* ('awake')	{NP Agent	_____	NP Patient/Theme }
		{NP Patient/Theme	_____ }	

Table 2 Verb-based syntactic transfer of transitive structures (Ingham 2020: 461)

According to Ingham (2020: 461), the OE verb *awacan* 'awake' was only used intransitively and gained a transitive argument structure in ME times (MED, 1250) through contact with the OF labile verb *esveiller*. The examples below aim to illustrate this.

OE, only intransitive:

(8) *He **awoc** of þam slæpe.*
he awoke of his sleep
(OED, Ælfric, Genesis ix. 24)

ME, intransitive and transitive:

(9) a. *On maregen ... he **awoc**.*
On morning ... he awoke.
(MED, c1175(?OE) HRood (Bod 343)2/15)
b. *Ich þouhte ... ðet ich wolde **awakien** ðe.*
I thought ... that I wanted awake you
(MED, a1250 Ancr.(Nero A.14)106/11)

OF, intransitive and transitive:

(10) a. ***Esveillierent** les andormiz*
(they)-woke the sleeping-ones
(TL, Perc. H 2077)
b. ***esvoiliez**, li juste, et si ne pechiez mie*
awake (ye) righteous and so not sin at-all
(TL, Job 334.)

Similarly, the OE verb *cleofan* 'split' was only used transitively and acquired an

intransitive argument structure (MED, 1325) on the model of OF (see Table 3)[4]. For Ingham (2020), these native verbs became labile via language contact with OF during the stage of bilingualism in ME times. Overall, these instances of syntactic transfer and the wholesale copying of labile structures from OF verbs (cf. Table 1) increased the amount of labile verbs in ME.

OE	*cleofan* ('split')	{NP Agent	_____	NP Patient/Theme }
ME	*cleven* ('split')	{NP Agent	_____	NP Patient/Theme }
		{NP Patient/Theme	_____ }	
OF	*croler* ('shake')	{NP Agent	_____	NP Patient/Theme }
		{NP Patient/Theme	_____ }	

Table 3 Verb-based syntactic transfer of intransitive structures (Ingham 2020: 461, slightly modfied)

Again, some examples for OE, ME and OF illustrate this.

OE, transitive:

(11) *In miclum gefyrstum timber* ***cleofan***.
in great intervals timber split
(OED, Gerefa in Anglia (1886) vol. 9 261)

ME, transitive and intransitive:

(12) a. *Ful wel kan ich* ***cleuen*** *shides.*
fully well can I split logs
(c1300 Havelok (LdMisc 108)917)
b. *Þe temple* ***clef****, þe eorþe qwok, þe dede arisen.*
the temple split, the earth quaked, the dead arose
(c1390 Dial.Bern. & V.(2) (Vrn)127)

OF, transitive and intransitive:

(13) a. *Vant l'ore et li raim* ***crollent***
blows the-wind and the branches shake
(TL, Rom. u. Past. I 5.)

4 Ingham suggests OF *croler* which, according to one reviewer, does not seem to have the same semantic properties that would facilitate this kind of syntactic transfer. OF *fendre* 'split' might be a better example.

b. *Li mur* ***crollent*** *et la torz tranblet*
the walls shake and the tower trembles
(TL, Ch. lyon 6542)

Although for Ingham (2020: 462) contact influence of OF led to an increase of labile verbs in ME, he concludes saying "[h]owever, since some lability in these semantic domains was present in OE, the role of French influence was not to create new argument structure possibilities for the language as a whole, but to drive the expansion of lability in English. We have argued that this took place via replication of the argument structure properties of French verbs in the relevant semantic domains, both with loans from French and with native-origin verbs belonging to these domains." From a learnability point of view it is not clear what Ingham means: does it imply that in his opinion the event structure and linking rules of labile verbs were already present in the English grammar given that there is a set of labile verbs in OE? In this case OF verbs filled verb slots in the already existing OE event structure and the effect of contact would be quantitative (only) (see also Percillier et al. 2024, Trips, Struik & Kaltenbach 2026). Alternatively, grammatical replication could also imply that a new structural template entered ME via OF (see also Trips 2014). To critically evaluate these claims we conducted a corpus-based quantitative study that we will lay out in the next section.

3.2 Corpus study: distributional patterns in Middle English

To ensure transparency and methodological rigor, we provide a detailed account of the corpus-based study[5], which serves as the historical linguistic foundation for our analysis.

Regarding corpus selection and temporal division, we conducted a study based on corpus data from the Penn Parsed Historical corpora of English to critically investigate Ingham's (2020) and GGI's (2023) studies. We used the three verb-lemmatised corpora available for the ME period (cf. Percillier et al. 2024 for details on the processing of the data): The *Penn-Helsinki Parsed Corpus of Middle English* (PPCME2, Kroch & Taylor 2000, release 4), the *Parsed Linguistic Atlas of Early Middle English* (PLAEME, Truswell et al. 2018), and the *Parsed Corpus of Middle English Poetry* (PCMEP, Zimmermann 2014). We refer to them as the combined ME corpora and refer to the periods that are pre-defined in the corpora as M1 (1150-1250), M2 (1250-1350), M3 (1350-1420), M4 (1420-1500). For the following more fine-grained analyses we also define more fine-grained periods of ca. 50 years. Argument-structure counts

5 Originally presented at the 57th Annual Meeting of the Societas Linguistica Europaea, Helsinki, 21-24 August 2024.

were done per predicate (IP). Table 4 gives an overview of the predicates per period.

Period	Year Range	Annotated IPs	Percent
1	1150-1200	14776	5.23
2	1205-1250	44511	15.74
3	1263-1300	29543	10.45
4	1305-1350	46366	16.40
5	1352-1400	54746	19.37
6	1402-1450	51144	18.09
7	1452-1500	41617	14.72

Table 4 Corpus composition and periods. Counts of annotated predicates (IPs) across 50-year periods in the combined ME corpora.

The process of extracting labile verbs was conducted in two stages. First, we identified all verb lemmas that occurred in either transitive or intransitive active constructions. To ensure statistical reliability and minimize the impact of hapaxes or transcription idiosyncrasies, we applied a frequency floor: only verbs with a cumulative token frequency of at least five occurrences within a given sub-period were retained for analysis.

Following the automated extraction, we applied two quantitative criteria to evaluate the distributional bias of each lemma.

i. Intransitivity Ratio (R_{intr}): This is the ratio of intransitive occurrences to the total number of transitive and intransitive tokens: $n_{intr}/(n_{intr} + n_{trans})$. A value of 0.5 represents a perfectly unbiased labile verb.

ii. Lability Score (L_{score}): To quantify the "distance" from a single-frame bias, we calculated a score on a scale of 0 to 1 using the formula: $1 - 2 \times |R_{intr} - 0.5|$.

Under this metric, a perfectly labile verb (50/50 split, $R_{intr} = 0.5$) receives an L_{score} of 1.0. For the purposes of our initial classification, we defined a verb as potentially labile if its lability score was ≥ 0.4, which corresponds to the minority frame accounting for at least 20% of the total tokens (i.e., R_{intr} between 0.2 and 0.8).

In terms of manual refinement and semantic classification, while the automated extraction provided a broad view of transitivity alternations, it inevitably included "noise" such as unergative verbs, object-drop, and reflexive constructions. To isolate the causative–anticausative alternation, we conducted a manual verification of the results. In the first stage of this research

[6], this manual cleaning was restricted to change-of-state (CoS) verbs. For the current article, this manual analysis was extended to include Change-of-Location (motion) verbs, following the semantic classes identified by GGI.

Concerning etymological information, we used the annotation of verbs for their origin from the BASICS project (e = english, f = french; for details see Percillier et al. 2024). This allowed us to calculate the proportion of French-origin verbs within the labile lexicon across each period. As shown in our previous results, the proportion of French verbs among labile types increases significantly toward the end of the ME period.

bin	period	etym	labile	trans	intr	mixed non-labile	labile ratio [%]
1	-1200	e	28	302	128	141	4.67
1	-1200	f	1	15	13	5	2.94
2	-1250	e	72	541	343	421	5.23
2	-1250	f	3	78	34	27	2.11
3	-1300	e	44	320	253	257	5.03
3	-1300	f	4	125	68	43	1.67
4	-1350	e	57	356	192	284	6.41
4	-1350	f	6	167	69	76	1.89
5	-1400	e	38	258	151	206	5.82
5	-1400	f	11	256	80	132	2.30
6	-1450	e	35	242	125	183	5.98
6	-1450	f	12	279	81	96	2.56
7	-1500	e	32	211	124	174	5.91
7	-1500	f	11	242	84	96	2.54

Table 5 Verb types per origin and period. "mixed non-labile" denotes verbs occurring in both frames but failing the lability threshold.

Figure 1 visualizes these findings. While the absolute count of confirmed labile verbs remains relatively low, the proportion of French-origin verbs within that set increases toward the end of the ME period.

The data suggest that the overall rate of lability is relatively constant and the influx of French verbs is not over-proportional to their general presence in the vocabulary. This supports our hypothesis that lability is an epiphenomenon of lexical composition rather than a structural shift driven by contact.

An explanation for the stable lability rate comes from cases where French labile verbs were copied into ME but lost their labile status as their event structure was narrowed. For example, OF *vuidier* ('to empty' or 'to become empty') was copied as ME *voyden*, but lost its anticausative argument structure except in restricted senses like the vacancy of an office. See the difference between

6 presented at the SLE, Helsinki, 2024.

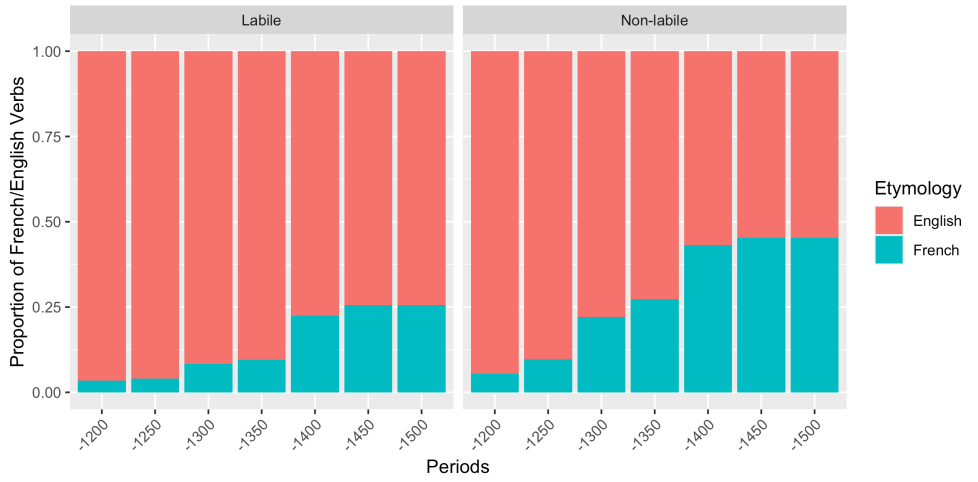

Figure 1 Relative proportion of English and French labile and non-labile verbs by period.

examples (14)–(15).

(14) a. *Par maintes fols* ***widas*** *les cofres.*
by many times emptied-you the chests
(Mir. de S. Eloi, p. 124, Peigné)

b. *Leonet de vif sont* ***wida.***
lion-cub of life was emptied
(Pastoralet ms. Brux. 11064, f* 62. v.)

(15) a. *Lat* ***voyden*** *al this hous in curteis wise.*
let clear all this house in courteous manner
(MED, (c1395) Chaucer CT.Mch.(Manly-Rickert)E.1815)

b. *He shold ... be promoted to the next bisshopriche that shall* ***voyde*** *wtin this reaume.*
he should ... be promoted to the next bishopric that shall become-vacant within this realm
(MED, (1454) Proc.Privy C.6.168)

In PDE, the verb *void* is used either as a transitive motion verb with the meaning 'leave' (16-a) or as a causative transitive in the formal sense of 'invalidate' (16-b):

(16) a. *She **voided** the lavatory, and ran cold water in the basin.*
(BNC, Dark dance. Lee, Tanith. London: Warner Books, 1993)
b. *The Harter Act **voided** any bill of lading clause that sought to relieve the carrier ...*
(BNC, Evolution and present state of the ocean bill of lading. Kozolchyk, Boris.)

The ME "counterpart" of Germanic origin is the verb *empten* ('to empty' < OE *ǣmtian*) that occurs only in a transitive causative structure:

(17) *And **empten** also grete and heuy purses.*
and empty also great and heavy purses
(MED, (c1395) Chaucer CT.CY.(Manly-Rickert)G.1404)

In PDE, *empty* is a labile verb with both transitive causative and intransitive anticausative argument structures, as shown in (18). In these cases, the native Germanic verb *empty* becomes labile after the ME period without any influence from OF.

(18) a. *He'd opened the bag and **emptied** the purse.*
b. *In the 19th century we were the world's Dubai, buying up art at eye-watering prices. In the 20th century the purse **emptied**.*
(*The Guardian*, Charlotte Higgins, "Who owns the Parthenon marbles?", January 15th, 2023)

There are also cases that show that OF non-labile verbs did not influence the argument structure of already existing labile verbs. One such case is OF *curer* 'to cure, restore to health' that occurs as a transitive causative verb:

(19) *Et Dieus Jesucris le **cura***
and God Jesus Christ him cured
(Sone 4633)

The ME verb *curen* robustly shows the same causative transitive structure in (20-a). There is one instance of the verb with an omitted object in the MED, see (20-b).

(20) a. *He **cured** Naaman, þe prince of Surre', fro seknesse of*
he cured Naaman, the prince of Syria from sickness and
lepre.
leprosy
(PPCME2, CMCAPCHR, 34.69)
b. *If he* [*a surgeon*] *schulde cure wiþoute knowleche þerof*
if he should cure without knowledge therof

[*of the cause*].

(MED, ?c1425 *Chauliac(2) (Paris angl.25)5a/b)

ME *helen* 'to cure, heal' of Germanic origin (< OE *hælan*) retained its labile property and occurs with both argument structures:

(21) a. *he tol=de is fader fore Hov a Naddre him hurte sore*
he told his father before how an adder him hurt severely
And hov Iesus ***hel=de*** *him*
and how Jesus healed him
(PLAEME, LAUD108AINFANCY.1154)

b. *Þe seste is þe wunde þt eauer wurseð on hond & strengere*
the sixth is the wound that ever worsens in hand and stronger
is to ***healen***
is to heal
(PPCME2, CMANCRIW-1,II.242.3517)

3.3 Conclusion: from history to learnability

These findings indicate that while OF provided a significant amount of labile verbs, the rules governing lability were already established in the English grammar. The fact that many French verbs met these conditions exhibits an overlap or a matching of identical structures in both languages rather than structural innovation (see also Trips 2020). However, the nature of this alternation—involving same-form verbs in two distinct syntactic-semantic frames—poses a specific challenge to the learner.

To understand how the historical child navigated the "noise" of remaining Germanic morphological markers (analogical leveling) alongside this French influx, we must move beyond descriptive corpus counts. In Section 4, we establish a formal learnability baseline using modern child-language data, redefining our lability metrics in terms of the TSP (Yang 2005, 2016). This allows us to test whether learnability thresholds can explain why certain classes generalized lability while others remained restricted. We assume the that lability (and causative-anticausative alternations in general) is rule-governed and licensed and constrained by the event structure of verbs. Under this assumption, the rules governing lability were well established in the ME grammar, and have always existed in English. French verbs copied to ME behaved exactly as predicted by these rules, and the proportion of French verbs among labile verbs does not justify the claim that structural copying took place. Rather, what we see are quantitative effects.

4 Lability in child acquisition data

4.1 Learnability and the uniformitarian hypothesis

The transition of the causative-anticausative alternation from a morphologically marked system in OE to a predominantly labile one in ME is often attributed to the significant influx of OF verbs. As outlined in the previous chapter, Ingham (2020) and GGI (2023) have provided a nuanced historical and contact-linguistic account of this process, suggesting that while the change was successful in domains like change-of-state (CoS) and motion, it remained temporary or failed in other classes, such as 'destroy' verbs and psychological (psych) verbs. GGI argue that the presence or absence of inherited structural "bridgeheads" in OE determined the long-term success of these contact-induced innovations (p. 354).

In this section, we address the points that Ingham (2020) and GGI have left unanswered regarding the applicability of modern learnability models and Pinker's (1989) semantic criteria to historical data: GGI suggest that the semantic constraints observed in PDE, such as the requirement for a specific resulting state to license lability, may not be immutable universal rules but rather diachronically emergent properties. Under this view, ME learners may have temporarily "relaxed" these constraints under the pressure of intense bilingualism and the statistical weight of the OF input (2023: 364).

After discussing previous accounts of lability acquisition in Section 4.2, we will address GGI's call for "deeper research" by re-evaluating the contact hypothesis through the lens of formal learnability in Section 4.3. Our position is that the general principles of acquisition remain constant over time: children partition the lexicon into subclasses based on salient conceptual features (such as those identified by Pinker) and evaluate the productivity of structural rules using precise thresholds. We contend that the apparent "relaxation" of rules observed by GGI may be the result of quantitative shifts in the learner's input rather than a fundamental change in the underlying learning mechanism. Following the logic of linguistic uniformitarianism (Labov 1972, Walkden 2019), we assume that the cognitive strategies and mechanisms like the TSP (Yang 2005, 2016) employed by contemporary children—specifically their sensitivity to type frequency—are the same ones that drove the diachronic shifts in ME.

Central to this mechanism is the distinction between Manner and Result (Pinker 1989, Levin 1993). Learners identify the domain (N) of the labile rule by filtering for verbs that lexicalize a resulting state (e.g., *break*, *open*, *melt*) while excluding those that merely describe a manner of motion or force without a specified outcome (e.g., *hit*, *push*, *kick*) and those that specify a man-

ner within the ACT structure (cf. Section 2). Under the uniformitarian assumption, the ME learner utilized these same conceptual primitives to identify which native and French-copied verbs were eligible for the alternation, regardless of their surface morphological marking. The application of this hypothesis requires an explicit model of how learners identify and generalize linguistic patterns. In the following section, we outline the theoretical foundations of the acquisition of lability, focusing on how children utilize conceptual features to define the scope of grammatical rules before they are tested against mathematical thresholds.

4.2 Theoretical foundations of the acquisition of lability

Research into the acquisition of the causative-anticausative alternation suggests that children do not learn transitivity on a purely item-by-item basis but rather by identifying semantic clusters that license the rule. Early developmental research identified recurring patterns of overgeneralization as a hallmark of the acquisition process (Bowerman 1974, 1982). Children systematically produce illicit causative or anticausative forms (e.g., **The toy disappeared the ball*), suggesting the generalization of abstract hypotheses about event structure rather than the rote storage of verb-specific frames. These empirical findings led to a debate on the domain in which bootstrapping emerged. Pinker (1989) argued for a *semantic bootstrapping* approach, where conceptual representations of causation and agency guide the child's mapping of participants onto syntactic positions. In this model, broad productivity leads to novel productions, while retreat is driven by accumulating verb-specific evidence. Conversely, the *syntactic bootstrapping* hypothesis (Gleitman 1990) proposed that children exploit the syntactic environment to infer verb meaning. Experimental evidence confirmed that even two-year-olds use distributional cues—such as the presence of a direct object—to interpret transitivity-alternating verbs (Scott & Fisher 2009).

More recent work focuses on the integration of conceptual, distributional, and morphological information. Children's sensitivity to the temporal and causal structure of events appears to be a robust constraint on acquisition: Concerning tense-aspect phenomena in general, van Hout et al. (2017) provide evidence that pre-school age children acquire event-structure constraints, e.g. the Agent Control hypothesis (Demirdache & Martin 2015: 201): "Zero-result non-culminating construals require the predicate's external argument to be associated with 'agenthood' properties." Thus, children are sensitive to the expression of changes of state, result states (perfectivity) and event culmination.

With respect to languages with morphological marking of the alternation,

Fotiadou & Tsimpli (2010) investigate how Greek-speaking children acquire the active vs. non-active morphology of the causative-anticausative pattern. Using elicited production and comprehension tasks, they show that children as young as four are sensitive to morphological voice marking, but they often overgeneralize active forms in contexts where adults prefer the non-active. The findings suggest that while the presence of overt morphology facilitates early recognition of alternations, full adult-like mastery requires integrating morphological, semantic, and lexical constraints.

An important factor for the language pair at issue is the fact that the acquisition of the alternation is modulated by the morphosyntactic transparency of the target language. In languages with systematic anticausative or reflexive marking, such as Romance or Greek, children exploit these cues early to constrain the relationship between transitive and intransitive forms (Bowerman & Croft 2008). In contrast, the relative opacity of the English system—which relies heavily on lability—results in different production profiles and developmental trajectories.

A central problem remains: the *retreat* from overgeneralization (e.g., of causative uses of intransitive verbs such as **The toy disappeared the ball*). The TSP (Yang 2005, 2016) is especially well positioned to account for this. What follows is a brief excursus introducing the TSP.

Productivity, as conceived of under the TSP, is a property of each individual's grammar. It is crucially categorical: patterns, rules, or generalizations are either productive or they are not. This binary distinction has been conceptualized in many theories of grammar (Anderson 1973, Kiparsky 1973, Aronoff 1976, *i.a.*) in terms of elsewhere conditions requiring irregulars to be looked up or accessed in some way before defaulting to a regular pattern. Since exceptions (i.e., one-offs or unproductive patterns) are ubiquitous across language, learners must be able to acquire, and speakers must be able to represent, patterns despite them. The TSP defines the number of exceptions that a hypothesized generalization may tolerate beyond which the speaker should give up and resort to memorization. The full cognitive motivation and mathematical derivation for the TSP are beyond the scope of this introductory discussion, but see Yang (2005), Yang (2016, pp. 10, 144), and Yang (2018) for more information.

When applied to concrete data, the TSP provides a quantitative metric for determining whether some linguistic pattern is productive in a speaker's grammar. While it was originally developed with morphology in mind, it is independent of any particular level of the grammar (Payne & Yang 2023), or even the narrow language faculty. It has been applied to a wide range of phenomena across phonology, morphology, syntax, and semantics across En-

glish (Yang 2016, Pearl & Sprouse 2018, Irani 2019, Richter 2021, Belth, Payne, Beser, Kodner & Yang 2021), German (Yang 2016, Belth et al. 2021), Icelandic (Nowenstein, Sigurjónsdóttir, Yang, Ingason & Wallenberg 2020, Björnsdóttir 2021), Turkish (Belth 2024), Cree (Henke 2023), and Kîîtharaka (Kanampiu, Martin & Culbertson 2025), among others, as well as extra-linguistic number cognition (Chu, Cheung, Schneider, Sullivan & Barner 2020), and social convention learning (Guilbeault, Caplan & Yang 2026). Motivated by research on pattern learning in psycholinguistics, the TSP has been validated in a range of experimental studies spanning morphophonology, syntax, semantics, and extra-linguistic cognition (Schuler, Yang & Newport 2016, Schuler 2017, Koulaguina & Shi 2019, Chu et al. 2020, Emond & Shi 2020, Shi & Emond 2023, Li & Schuler 2023, Guilbeault et al. 2026). It has also found use in diachronic modeling of language change through acquisition, which we leverage in the next section of this paper.

The TSP is realized in two ways: as the Tolerance Principle (TP) and Sufficiency Principle (SP). Our paper relies on the SP, but since it derives from the TP, we introduce it first with the formal definition provided in (22). Core to the TP is the *tolerance threshold* θ_N, which is the number of exceptions below which a learner should acquire a rule or generalization. θ_N is a function of N, the number of lexical items in the scope of the generalization, and e is the number of exceptions known to the learner so far. It is defined as the number of known types that a generalization should apply to divided by the natural logarithm of the number of types.

(22) **Tolerance Principle**: If R is a productive rule applicable to N candidates, then the following relation holds between N and e, the number of exceptions that could but do not follow R:

$$e < \theta_N \text{ where } \theta_N := \frac{N}{\ln N}$$

The SP reframes the TP in terms of sparse attestation. Rather than asking if the learner has learned tolerably few exceptions to a hypothesized generalization, it asks whether the learner has received a sufficient number of generalization-following items to commit to it. If N has the same meaning as for the TP, and M is the number of items in the scope of the generalization which the learner knows to obey it, then $N - M$ is the number of items for which the learner is lacking evidence. In principle, each of these $N - M$ items could constitute exceptions to the generalization if they were attested, so the SP represents a worst-case calculation for the TP. The SP thus uses this number instead of e when calculating reliability. As shown in (23), the threshold

is otherwise defined in the same way (cf. Yang 2016: 177).[7]

(23) **Sufficiency Principle**: Let R be a generalization over N items, of which M items are attested to follow R. R can be extended to all N items if and only iff:

$$N - M < \theta_N \text{ where } \theta_N := \frac{N}{\ln N}$$

Importantly, as a learner's lexicon develops, the number of known exceptions or unattested forms relative to N will change. A pattern that was productive early in development may later be rendered unproductive, or vice-versa. This, then, provides a mechanism for the *retreat* from overgeneralization. Recent learning-theoretic accounts such as Irani (2019) frame this as a statistical learning problem. For Irani, the basis for learning the causative-anticausative alternation is a subclass of intransitive verbs that can also be used in transitive causative structures. In her study of CHILDES data, she evaluates the productivity of the rule: 'intransitive change-of-state verbs can be used transitively" applying the SP where N is the number of change-of-state verbs attested in the learner's vocabulary, and M is the number of these verbs also attested in transitive construction. She concludes that children learn the causative-anticausative alternation on the basis of positive evidence of verbs that occur both intransitively and transitively whereby the transitive causative verbs define the SP threshold (see also section 4.3 for our approach).

Given these theoretical constraints, we now evaluate the available signal in naturalistic input. If children utilize semantic classes as the reference set for the TSP, we should be able to observe whether modern child-directed speech provides a sufficiently clear signal to license a productive labile rule for the classes, similar to the hypotheses of Pinker (1989) and others.

4.3 Labile verbs in child acquisition data

To provide the ground for the learnability model, we first examine how modern children acquire the causative–anticausative alternation using the TSP. This exploration addresses a fundamental question for the ME transition: how much evidence does a learner require to generalize a labile rule? While our previous historical investigations focused primarily on CoS verbs, for the

7 For both the TP and SP, there is no difference between defining the threshold in terms of inequality ($\leq$) or strict inequality ($<$). Since θ is never an integer while e and $N - M$ always are, equality is impossible. It has been defined in both ways in the literature (compare Yang (2016) to Yang & Montrul (2017)). We have adopted strict inequality for both definitions for consistency and to emphasize this point.

current study, we extended the scope to include motion verbs—a class identified by GGI as a secondary domain of successful lability. This expansion allows us to compare the productivity of a lability rule across classes of different sizes, providing a more rigorous test for the consistency of the learning mechanism.

Before learners can evaluate the productivity of a class-wide rule, they must first determine which individual lemmas in the input are [+labile]. We explore this acquisition baseline using PDE data from the CHILDES database, extracting 3 million utterances from 55 North-American projects (MacWhinney 2000). A significant methodological challenge in analyzing child-directed speech (CDS) for lability is the nature of automated parser annotations. We used the UDPipe parser (Straka 2018) in the conversion pipeline described in Stein (2026). Syntactic parsers distinguish transitive uses from intransitive ones but do not sub-categorize the latter into "anticausatives" (where the object of the transitive corresponds to the subject of the intransitive, cf. Section 2) versus other intransitive constructions, such as unergatives or object-drop. We sub-classified intransitive structures utilizing a heuristic-based subclassification to identify true anticausatives (e.g., *The window broke*) rather than an underspecified "non-causative" category relying on the absence of the direct object alone. By filtering for instances where the subject of the intransitive corresponds to the potential patient of the transitive, we reduce the noise introduced by agentive (unergative) uses or object-drop.

We started the mapping process with the curated set of 102 ME verbs of change-of-state ($N = 78$, henceforth "CoS") and motion ($N = 24$) that were analysed as "labile", based on corpus analysis and dictionary information (cf. the list in Appendix C). To facilitate a merge with modern data, these historical lemmas were first expanded based on their modern translations. In instances where an ME lemma was associated with multiple synonymous translations (e.g., comma-separated entries in the curated table), they were separated into individual rows to ensure that no potential modern matches were overlooked. The resulting expanded historical list was merged with the lemmatized CHILDES dataset. This mapping process yielded a final set of 68 matched verbs (listed in Appendix C) for which syntactic data was extracted from the CHILDES subcorpora.

In Figure 2, the y-axis represents the difference $Ratio_{child} - Ratio_{adult}$, where the dotted zero line indicates identical ratios across both groups. A significant relationship exists between adult input frequency (on the x-axis, increasing from left to right) and the child's causative bias (Pearson's $r = 0.404$, $p = 0.0332$; Spearman's $\rho = 0.458$, $p = 0.014$). We apply a log-transformation to frequency to normalize the Zipfian distribution of the lexicon, ensuring

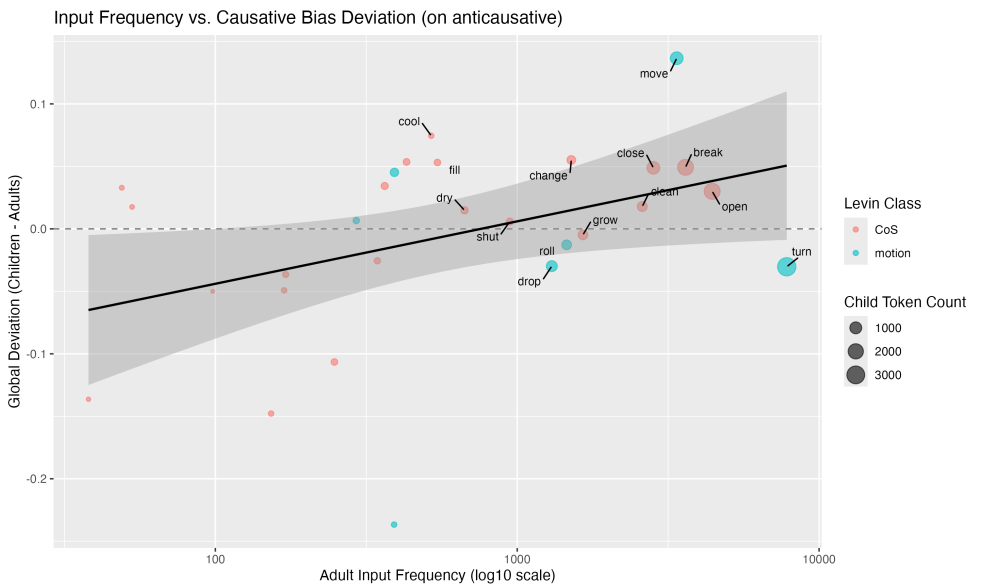

Figure 2 Comparison of causative bias in child-directed speech.

that the analysis is not skewed by high-frequency outliers like *break* or *move*.

The data reveal a clear developmental shift tied to input frequency: while lower-frequency verbs are associated with a greater "anticausative gap" (where children under-use the causative frame compared to adults), this gap effectively closes for high-frequency items. For core triggers such as *break*, *change*, and *move*, children exhibit causative ratios that match or slightly exceed those of their caregivers. This apparent over-use suggests two possible interpretations. On the one hand, high-frequency triggers may act as particularly robust causative anchors in the child's emerging grammar, leading to an over-generalization of the transitive frame for these specific items. Alternatively, this pattern may reflect genre-specific properties of the adult CDS baseline; if caregivers in a nursery setting under-use causative structures relative to broader PDE norms, the child's higher ratio may represent a convergence toward a more general linguistic target. While the real cause for the over-use is not relevant for our primarily historical argument, the significant correlation nevertheless confirms that input frequency is the primary driver in resolving the competition between structures at the lemma level before class-wide rules are generalized, regardless of which interpretation is

true.

We model lability learning with a two-step calculation: Before a learner can evaluate the productivity of a class-wide rule with the TSP, they must first determine which individual lemmas in the input are [+labile]. We model both stages:

Stage 1: Detection of labile verbs, using a fixed 10% detection limit. N = total tokens per lemma; e = minority frame tokens. A verb is categorized as [+labile] in the corpus if its minority frame (causative or anticausative) exceeds 10% of its total tokens. Verbs with fewer than 5 tokens are ignored ($N \geq 5$), as these are two rare for the 10% limit to cleanly apply and approach the limit of the mathematical applicability of the TSP.

Stage 2: Class and Global productivity, using the SP. N = total verb types per class; M = labile types identified in Stage 1. At this level, we evaluate whether the number of [+labile] types (M) identified in Stage 1 is sufficient to generalize a rule across a semantic class of size N. Following the SP, a rule (e.g., "CoS verbs are labile") is considered productive if M meets or exceeds the sufficiency threshold defined in (23). This calculation is applied both to the individual semantic classes and to a global conflation of all verbs of change to test the limits of the learner's partitioning.

In order to manage possible noise in the data and apply a stricter metric for lexical learning than simple attestation in CDS, we focus in Stage 1 on a 10% fixed detection limit for identifying labile verbs (that is, a given verb's minority construction must account for at least 10% of its total tokens), though we consider others as well. This 10% limit is a common heuristic in quantitative linguistics to distinguish systematic alternations from performance errors or idiosyncratic noise. The results of this stage of analysis show that high-frequency triggers like *break*, *turn*, and *grow* clear the 10% threshold easily (Figure 3). Interestingly, under the strict "anticausative" filter, many motion verbs—including *move*, *roll*, and *slide*—fail to reach even the 10% mark in CDS, appearing almost exclusively in transitive causative contexts.

At the second stage, we evaluate class-wide productivity by applying the Sufficiency Principle over the 44 verbs determined to be labile in our analysis. The results, summarized in Figure 4, provide a striking contrast between the semantic classes. The bars represent the number of observed labile types (M), while the markers indicate the sufficiency target required to license a productive rule.

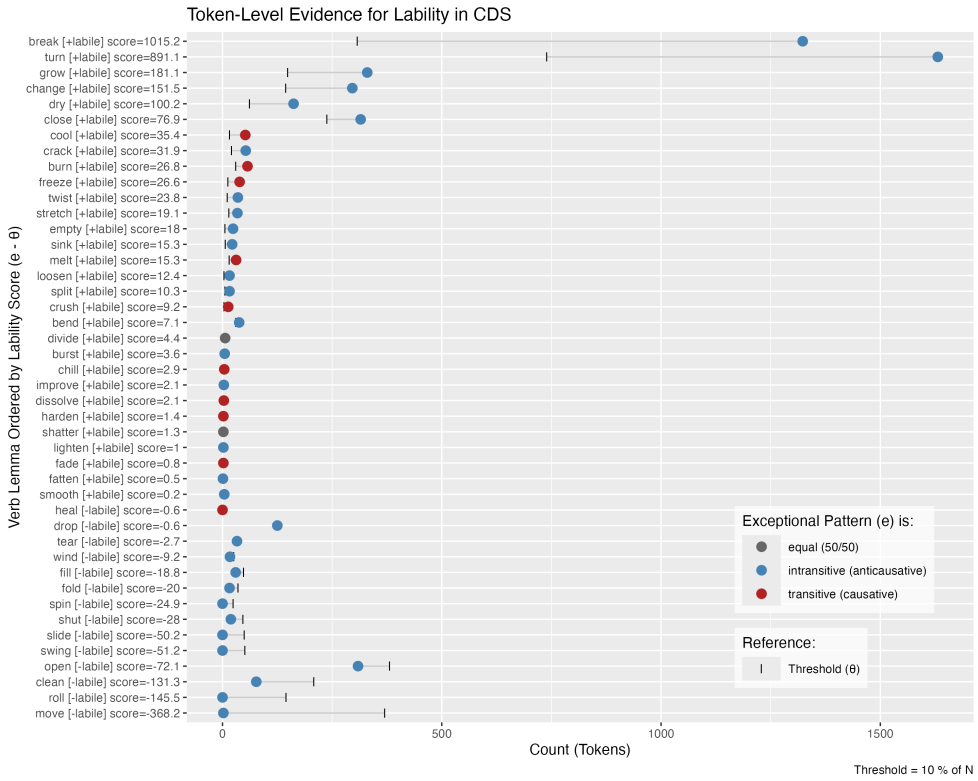

Figure 3 Token-level evidence for lability. Values beyond the 10% detection limit (bars) identify confirmed labile types (M).

With $N = 35$ and $M = 28$ labile types, the CoS class is productive under the SP ($N - M = 7 < \theta_{35} \approx 9.8$). The motion class, with $N = 9$ but only $M = 2$ labile types, fails the SP threshold ($N - M = 7 < \theta_9 \approx 4.1$). The motion lability rule is not productive. This suggests that while lability is a robust, generalized rule for CoS verbs in PDE (at least, in the genre of CDS), it remains a lexicalized, item-by-item property for motion verbs.

Combining the CoS and motion classes, a global rule with $N = 44$ fails, because its observed $M = 30$ is less than the required sufficiency target for $N = 44$ ($\theta \approx 11.6$ exceptions tolerated, meaning M must be ≥ 32.4). This confirms the "partitioning advantage" for the learner: by treating CoS and motion as distinct domains, learners are able to generalize a productive rule for the CoS class, whereas a global rule would be unproductive due to the

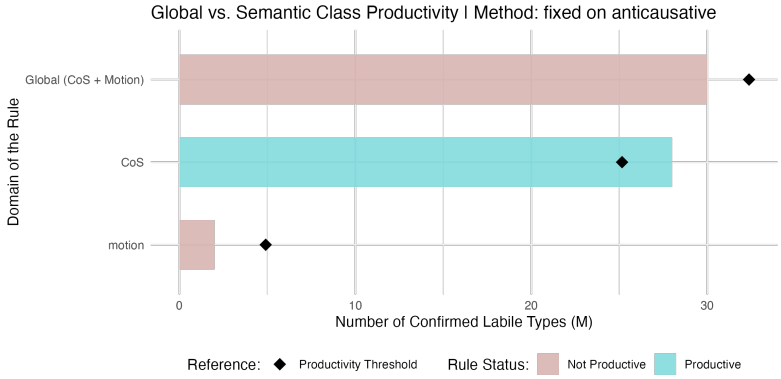

Figure 4 Aggregated productivity for semantic classes, separately and conflated. Bars represent the number of labile types; diamond markers indicate the productivity target.

noise of the combined dataset.

These outcomes depend on the chosen detection limit. If a more stringent 20% detection limit were chosen, core triggers like *close* or *burn* would be lost to the learner as exceptions. The CoS rule fails because the number of exceptions exceeds the tolerance threshold ($N - M = 15 > \theta_{35} \approx 9.8$). Conversely, a lower 5% limit renders both the CoS and global rules productive, though we contend that this is the result of misidentifying infrequent object-drop productions as evidence for a grammatical rule. Finally, the Motion rule would require a 0% detection limit to achieve productivity. Taking all the thresholds into account, we find that the CoS pattern has the most robust evidence behind it for productivity. The Global pattern is achievable with a more permissive threshold, though this may be due to the increased chance for misclassifying verbs in the input. There is no reasonable support for the Motion rule.

The results from this CHILDES exploration establish that while the CoS rule is productive today, it is mathematically sensitive to the ratio of labile types in the input. In the following section, we return to the ME data to test the hypothesis that the French lexical influx acted as the necessary statistical catalyst to resolve this bottleneck. Here, we reintegrate the etymological and morphological distinctions of Section 3 to simulate how contact-induced lexical expansion interacts with internal morphological leveling.

5 A Diachronic Simulation

In this section, we present a diachronic simulation for the development of lability in Middle English. Adopting the widely discussed causal relationship between children's innovations during language acquisition and population-level language change, we leverage the quantitative TSP here, on ME data, to gain a deeper understanding of the effects of morphological leveling and French copying on lability.

One important property of the TSP is that learners will adjust their decisions about productivity as they learn new vocabulary and their tolerance thresholds shift accordingly. This is what accounts for the advance and retreat of overgeneralizations during development (Yang 2016), including for the causative alternation (Irani 2019), and also plays a role in accounting for child-driven actuation of grammar change (Kodner 2023a). We track this in our simulations by recalculating the TSP repeatedly as simulated learner vocabularies increase. Ideally, we would estimate learner vocabularies from a corpus of child-directed speech (CDS) as we did in the previous section, but this presents an obvious problem for a simulation of Middle English learners, as there is a complete lack of CDS corpora available for the language. This challenge is not unique to this language, or even to historical languages in general, as there is no CDS available for the vast majority of modern languages either. Fortunately, it has been argued that historical corpora provide suitable type frequency estimates for applying the TSP (Kodner 2019a), which we now describe.

There are two properties of acquisition and early linguistic input that facilitate historical corpus investigation. First, it is type frequency, not token frequency, over which a pattern is expressed that drives productivity learning. Results from a wide range of often diametrically opposed research programs agree on this point (Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016). Consequently, it does not actually matter which specific words are present in the lexicon (derived from CDS or otherwise) as long as the lexicon is of approximately the correct size and contains roughly the correct proportion of words that express a pattern. In this regard, the TSP is especially forgiving, since the actual number of exceptions does not matter, as long as it falls on the right side of the threshold. Differing vocabulary across genres, be they CDS or historical texts, may matter less than one would expect.

Second, the relative age at which words are acquired correlates with their token frequencies (Palermo & Eberhart 1968, Goodman, Dale & Li 2008, Swingley & Humphrey 2018, Braginsky, Yurovsky, Marchman & Frank 2019). This makes sense because high frequency items are more likely to be attested

to learners early on in their development, since they have more opportunities to appear. Early learner vocabularies are quite small. Across languages that have been investigated, the typical learner knows only a few hundred to a thousand lemma by around age three when most of morphology acquisition has taken place (Fenson, Dale, Reznick, Bates, Thal, Pethick, Tomasello, Mervis & Stiles 1994, Hart & Risley 1995, 2003, Kern 2003, Szagun, Steinbrink, Franik & Stumper 2006) and only a third to half of those are verbs (Bornstein, Cote, Maital, Painter, Park, Pascual, Pêcheux, Ruel, Venuti & Vyt 2004). We can estimate a learner's vocabulary by sampling, directly or probabilistically, several hundred of the most frequent words from a corpus. Then, since words, regardless of corpus genre, tend to distribute according to a Zipfian distribution (Zipf 1936, Jelinek 1997, Chan 2008, Yang 2013), this process of sampling high-frequency words has the effect of removing the long tail where most corpus-specfic and genre-specific vocabulary lies. Given these two properties, the few hundred most frequent types in a corpus of CDS makes a good vocabulary estimate for a child learner (Nagy & Anderson 1984, Yang 2016).

Kodner (2019a) shows that since applying a frequency threshold removes mostly low-frequency document-specific or genre-specific vocabulary, lexical distributions tend to converge across genres when only a child lexicon-sized number of high-frequency words is sampled from a corpus. Thus, we can take the high-frequency portion of the sample to approximate the lexicon of a "typical" child learner, and historical and other non-child-directed corpora can often be substituted for CDS when the same processing steps are applied. While lexicons derived from CDS and non-CDS corpora do differ, the high frequency items from these corpora are roughly as similar as those derived from different CDS corpora are to one another. Trips & Rainsford (2022) provides additional validation of this reasoning, including discussion of the conceptual implications for treating "panchronic" historical data drawn from periods lasting potentially many years, to synchronic learners and grammars. Given these findings, we are in a strong position to apply the TSP to our ME data in order to investigate lability in the historic language.

5.1 Simulations

Our first simulation provides an initial assessment of the impact of French verbs in the lexicon on the learnability of lability in ME using the Sufficiency Principle and the lexicon of CoS verbs extracted from our data. This establishes a few points. First, it provides a "synchronic" test for the learnability of productivity for ME learners. This method can establish whether the productive pattern is actually learnable from the data we have at hand. Doing so

lends credibility to our substitution of historical text for CDS. Second, it establishes something about the pattern's learning trajectory by indicating roughly what proportion of ME's CoS and/or Motion verbs would be learned before lability is rendered productive. Third, it establishes our simulation methodology, in which we perform frequency-weighted sampling on our data set to approximate the learning trajectories of a "typical" population of children. Fourth, we manipulate the data in order to assess the role that French verbs in the ME lexicon play on the productive outcome for this generalization. It provides a way of quantitatively assessing the contact hypothesis.

To simulate a learner's vocabulary growth and exposure to verbs in both transitive and intransitive frames, we perform frequency-weighted sampling. This procedure, outlined below, is performed 10,000 times to model the lexical development of 10,000 children. Our ME data set provides token counts for each verb as it appears in both transitive and intransitive frames. We take these counts to approximate the distribution of these verbs in these frames in a typical learner's input. The full set of tokens is shuffled for each learner and presented to the learner in the resulting order. The learner then tracks which lemmas have been encountered, and in which frames. Sampling ends once the learner has encountered all of the tokens in the corpus. Further filters were applied to bring this simulation in line with the analyses of modern CDS, namely a 5-token minimum threshold to count a verb in the lexicon and a 10% detection limit for lability. For each simulated child N is calculated as the number of unique verb types that the child has encountered up to that point, and e is the number of those types which are not evidently labile (i.e., they fail to meet the detection limit at that point in the simulation).

This methodology is mathematically equivalent to filtering infrequent verbs from the data, then sampling from the set of lemmas without replacement, weighted by frequency in transitive and intransitive frames. Such frequency-weighted sampling algorithms have become standard in modeling learning trajectories with the TSP (Belth et al. 2021, Kodner, Payne, Khalifa & Liu 2025). Sampling in this way improves upon the simple frequency thresholding of Kodner (2019b), because each sample can be used to represent a different learner's input.

In order to simulate learning trajectories, learners recalculate the productivity of lability whenever a new verb type is encountered. The trajectories are averaged across simulated learners, so they can be interpreted as the likelihood of a child acquiring a productive generalization at a given relevant vocabulary size or the proportion of children expected to acquire the generalization at that size. A high or low value indicates that the generalization is robustly productive or unproductive across learners, and middle values

indicate substantial inter-learner variation subject to the specifics of their individual inputs.

Simulations were performed for the CoS (up to $N = 81$), Motion (up to $N = 30$), and combined Global (up to $N = 111$) generalizations. The learning trajectories averaged across all simulated children are displayed as the black lines on the left side of Figure 5. For CoS verbs, productivity is almost immediately acquired by nearly all learners, as depicted by the near-ceiling value. Early productivity opens the door for overgeneralizations, so this may be consistent with the types of overgeneralizations observed in PDE (cf. Bowerman & Croft 2008). Motion verbs show a different U-shaped averaged pattern in which all learners initially acquire a productive generalization, then about a third of them fall out of productivity before eventually regaining it at the very end of the simulation. The majority retain the generalization throughout learning, as shown by the U-shape never dipping below 50%. The Global generalization splits the difference between CoS and Motion, showing rapid generalization followed by a subtle dip affecting up to about a quarter of learners about a quarter of the way through acquiring the lexicon. This result follows from the definition of the Global generalization, which combines CoS and Motion verbs.[8]

This overall result differs from our synchronic study, which only found productivity for CoS verbs. There are two possible explanations for this. First, it may indicate substantive differences in the lexicons of PDE and ME, which are effectively different languages, or they may be the result of meaningful genre differences contra our assumptions. Kodner (2019b) opens the door for interoperability of CDS and historical corpora, but it is not to be taken as a universal license for using historical corpora in this way. Second, our analysis of PDE models only a single child. Learners over our ME data who have acquired an equivalently sized lexicon show variation in which generalizations they have acquired, and our PDE child may correspond to learners who have not acquired all three generalizations. Our ME simulation lends support to this second possibility: By varying the detection limit, we found that on modern CDS data, the CoS rule was most robust, followed by Global, followed by Motion. We observe that same ranking on the ME data, as manifested in the depth of the U-shaped learning trajectories. Furthermore, the depth of the U-shaped learning trajectories suggests that many learners, nearly 50% in the case of the Motion rule, should fail to generalize during a large period of their development.

8 The jagged lines which appear in all simulations are a consequence of the threshold nature of the TSP, and are not particularly meaningful. A peak occurs whenever θ exceeds a whole number, since that increases the tolerable $N - M$ by one.

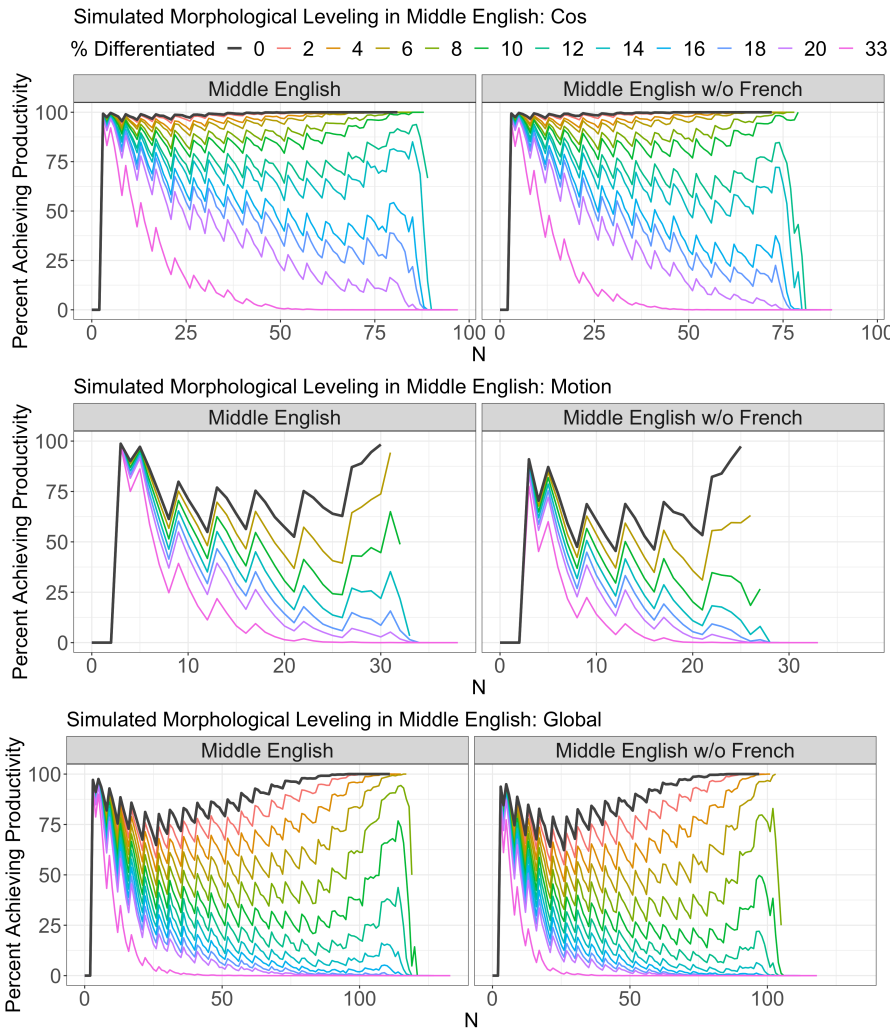

Figure 5 Averaged learning trajectories for 10,000 simulated Middle English learners by condition, with simulated morphological differentiation. The black lines indicate learning trajectories on the actual Middle English data and Middle English excluding French copy verbs. The colored lines indicate average learning trajectories when morphological differentiation is reimposed on different proportions of verbs. Beyond 30% differentiation, no learners acquire labile generalizations.

With these basic results established, we present a slight modification to our initial simulation which evaluates whether the presence of French vocabulary in Middle English was "synchronically" critical for learning the lability

generalization. To do this, we reran the same simulations with all words of French origin removed from sampling (leaving CoS $N = 72$, Motion $N = 25$, combined Global $N = 97$). If removing French-derived verbs delayed or prevented learners from reliably acquiring productive generalizations, this would point to a crucial role for these verbs in facilitating lability.

However, we find that simply removing the French verbs have little effect on the learning curves, as visualized as the black lines on the right side of Figure 5. For all three generalizations, the average learning curves over ME data with French verbs removed line up closely with the full data set ME curves. From these results simulating typical learning trajectories in Middle English, we can conclude that the presence of French copy verbs was not necessary for Middle English learners to acquire the labile generalization. In fact, they made little difference at all by the time of the periods from which this data is collected, since English-derived verbs were already in labile use.

The simulations presented so far, while modeling the past, are effectively synchronic in nature. We transition to diachronic modeling in an evaluation of the quantitative impact of morphological leveling on the feasibility of lability in ME. This time, a percentage of native labile verb types is chosen to be morphologically differentiated in their intransitive and transitive frames, reflecting an earlier state in the language when fewer verbs had been leveled (for example, the distinct OE causative pairings like intransitive *sincan* 'sink' versus transitive *sencan* 'cause to sink,' cf. (4)). This process increases the potential N and e for each generalization, as the intransitive and transitive verbs count as distinct types, neither of which can provide evidence for lability. Since French copy verbs were never morphologically differentiated in English, they were exempted from differentiation. This provides a more subtle avenue for French verbs to facilitate the development of lability in the language.

Two patterns emerged from the simulations, as shown as the colored lines in Figure 5. First, for all three generalizations, whether or not French copy verbs are included, the likelihood of generalization drops off precipitously as the proportion of potentially-labile verbs which are morphologically differentiated increases. This is indicated by the colored lines in the figure. If 18% or more relevant verbs are differentiated, then no generalization is productive over the full set. Most learners initially posit productive lability, but the generalization gradually becomes untenable as more and more non-alternating verbs are acquired. This points to morphological leveling as a necessary catalyst for the acquisition of the labile generalizations: as more and more verbs were leveled, productive lability became more and more tenable. Furthermore, this predicts that *most* morphologically differentiated potentially labile

verbs must have been leveled for the generalization to have emerged. CoS is ultimately learnable by 100% of learners over the full data when 10% or fewer of the verbs are differentiated. The Motion and Global generalizations have stricter requirements, with Motion requiring 0%, and Global requiring 6% or less for all learners to converge on productivity. Both generalizations show variation when fewer than 10% of verbs are differentiated.

The second pattern, observable by comparing same-color lines in the left and right columns, is that the presence of French copy verbs in the lexicon, which always have the potential to alternate, renders labile generalizations slightly more robust to morphological differentiation. For example, nearly all learners are expected to acquire the CoS labile generalization when 10% of verbs are morphologically differentiated when the full Middle English lexicon is considered, but none are expected to do so in the end when French verbs are excluded. Similarly, curves for the full ME lexicon are elevated over corresponding curves with French verbs removed, indicating that at any given time, a larger portion of learners have acquired the productive generalization. This points to just a small acceleratory or facilitatory effect of French influence when combined with morphological leveling on the feasibility of the labile generalization. It is important to note, though, that this account, which is purely quantitative and does not meaningfully track lexical identity, is not equipped to evaluate the proposal that structures were copied wholesale from OF verbs to OE verbs with the same meaning as proposed by Ingham (2020) and discussed in Section 3.

5.2 Putting together an account for actuation

Modeling of individual acquisition in the past, such as that described thus far, provides an account for the innovation of novel generalizations. Some learners, who happen to acquire a lexicon with favorable distributional properties, may propose a (perhaps transient) innovation. But a satisfactory account for actuation extends beyond simple innovation, since any innovative piece of grammar must make the leap from the individual to the speech community in a process called actuation (Labov, Yaeger & Steiner 1972). While the exact mechanisms by which early learner innovations actuate language change remain mysterious, Kodner (2023a) and Dolatian & Kodner (in press) provide a model for this type of productivity-driven actuation. Drawing on the experimental and observational literature on language acquisition and variation and peer-to-peer transmission (Hudson Kam & Newport 2005, Nardy, Chevrot & Barbu 2014, Austin, Schuler, Furlong & Newport 2022), they emphasize that actuation need not entail that learners ultimately stick to their own innovations. This is because an innovation can take on a life of its own

as it is transmitted peer-to-peer, whether or not its original innovator eventually grows out of it.

With that in mind, it is not only the typical learner's final state, but rather the typical learning trajectory, that is most relevant for actuation. Innovative learners adopting lability should sometimes produce innovative constructions in which verbs appear in frames disallowed by their predecessors' conservative grammars. Given the sparsity of such input, other learners may either fail to recognize these constructions as innovative and adopt them, or they may identify them but adopt them in the course of developing sociolinguistic competence. These learners then provide positive evidence for the innovative grammar, so subsequent learners need not re-innovate them. The more learners who may adopt a novel generalization at some point, as well as the longer span over which learners are likely to retain the innovation, the more opportunities there are for this kind of transmission.

Returning to our simulations, even at the maximum morphological differentiation, most learners still pass through a phase of early productivity for each of the three generalizations.As more verbs are leveled (percent differentiated decreases), the average learning curve consistently raises, so that more and more learners pass through an extended phase of productive lability. In this range, actuation of the innovative grammar is increasingly possible through peer-to-peer interaction. Between 8% and 12% differentiation, there is substantial inter-learner variation in whether or not the labile grammar is ultimately acquired from syntactically conservative input, and below 8%, it is inevitable. Comparing the simulations with and without French copy verbs, it can be observed that the presence of these verbs (the left column) shifts the average learning curves upward (compared to the right column). This indicates that a larger proportion of learners entertain productive lability at some point during their development because of these verbs regardless of the rate of morphological leveling. These learners would be in a position to actuate the change.

Grounding the particulars of the simulation's numbers is challenging because the time resolution of our Middle English data set does not allow us to recalculate productivity over successively more morphologically leveled lexicons. We can, however, make a quantitative estimate from García García (2019), which provides a list CoS and Motion verbs which underwent morphological leveling between Old and Middle English. Out of 86 verbs, 29 (33%) retained morphological differentiation. Since this wordlist is effectively a dictionary and not drawn from actual use directly, this 33% number is not directly comparable to our simulations, but we have included it as a curve in the figures for completeness. It does, however, confirm that there was dra-

matic morphological leveling of these verbs in the relevant time period. It also suggests that the real proportion of morphological differentation at the time of the change is on the higher end of what we have simulated, pointing to a scenario where actuation occurred while the input evidence was not overwhelming for lability.

In sum, these TSP simulations lay out a scenario in which morphological leveling was the primary enabling factor for the actuation of productive lability, in line with García García (2019), while the presence of French verbs played a relatively minor but facilitatory role. Thus, the "relaxation" of semantic constraints noted by GGI is reinterpreted here not as a change in cognitive rules, but as an inevitable result of the shifting lexical distribution in the input.

6 Conclusions

The primary objective of this study has been to determine whether the considerable increase in labile (causative–anticausative) alternations in ME represented a fundamental shift in the English grammar or a predictable result of lexical dynamics. By integrating the contact-linguistic hypotheses of Ingham (2020) and García García & Ingham (2023) with the formal constraints of the Tolerance and Sufficiency Principle(s), we have proposed a model where language change is driven by the interaction of internal morphological leveling (from OE to ME) and external lexical influx (from contact with OF).

From a cognitive perspective, our findings suggest that the rise of lability was not a relaxation of semantic constraints (which GGI proposed as a tentative explanation), but rather a mathematical consequence triggered by changes in the input. In Section 4, we demonstrated that modern child learners' data can serve to establish an acquisitional baseline: if learners can successfully extract a productive rule for change-of-state (CoS) verbs from a noisy input (regarding the choice and the bias of the minority constructions of alternating verbs), the ME learner—operating under the same uniformitarian constraints—would have done the same once the critical mass of triggers was reached.

While we do not reject a role for OF contact emphasized by Ingham (2020), our results refine the nature of its influence. Rather than OF replicating a new structural template, the influx of French verbs acted as a support in the acquisition of the English grammar. The simulations in Section 5 demonstrate that French verbs increase the proportion of simulated learners who ultimately acquire productive generalizations.

A central contribution of this paper is the demonstration of a new methodological pathway for historical linguistics: the combination of historical cor-

pus linguistics with cognitively motivated computational modeling. While "standard" historical corpus linguistics provides the phenomenon, i.e., the rising frequencies of same-form verbs and the loss of morphological markers, it cannot explain alone the timing or the success of grammatical change. By applying the TSP, we provide the explanation of why the changes occurred: it allows us to identify the specific mathematical "tipping point" where a collection of individual lexical items becomes a generalized rule. This synthesis has allowed us in Section 5 to model how much lexical contact is required to trigger a systemic change in grammar.

The anticausative heuristic established in Section 3 provides a good starting point for further comparative research. A natural next step is to expand this analysis to languages with varying degrees of morphological marking. For example, German has retained significantly more affixes for transitivity and is expected to provide a higher number of exceptions to a labile rule. Testing the TSP scores of German verbs could reveal if its lower rate of lability (in the narrow sense) is a direct consequence of a higher mathematical barrier to productivity. Among the Romance languages, French and Italian, to different degrees, utilize strategies like the reflexive marker to distinguish causative from anticausative. A TSP-based comparison would clarify whether these explicit markers lower the "noise" in the input, allowing for faster rule generalization.

In conclusion, the rise of the English labile verb was a result of internal and external forces. Including the child learner as governed by cognitive principles of learnability, we were able to reconcile the qualitative semantic classes of Pinker (1989) with the quantitative realities of the ME historical data. We could do this only by applying cognitively motivated computational modeling to historical data to reveal how the historical child acquired and learned labile verbs.

A Data supplement

The appendix lists our resources (lists of verbs, quantitative information etc.) if they are not too large. For data tables and other materials unsuitable to be published here, please consult the data supplement provided at Open Science Framework (OSF), https://doi.org/10.17605/OSF.IO/GSPV2. Code for the simulations is available on Github, https://github.com/jkodner05/Lability_JHS2026.

B Materials for Section 2

B.1 Verbs extracted from the combined Middle English corpus

Cf. Section 3.2: The extraction of labile verbs followed a two-step procedure combining automatic filtering with manual philological verification. Initially, all lexical heads in the combined ME corpus were retrieved if they appeared in both transitive and intransitive constructions within active clauses. From this list, we calculated an intransitivity ratio and a lability score to identify verbs with a significant presence in both patterns.

To filter out noise from object-drop or unrelated polysemy, the resulting list was manually reviewed. Verbs were assigned to semantic classes—primarily change of state (CoS), motion, psych verbs, and aspectual verbs—and only those confirmed to exhibit a true causative-anticausative alternation were included in the final qualitative analysis (indicated by the verified flag in the data supplement). The final dataset provides a comprehensive overview of 142 verified labile verbs across all verb classes, their etymological origin, and their frequency in the combined ME corpus.

CoS verbs were used in the historical study (Section 3.2). CoS and motion verbs were used in the acquisitional study (Section 4.3 and Appendix C).

- 86 Change-of-state verbs: ablenden (‘blind, dazzle’), adrenchen (‘drown’), aliggen (‘alleviate, lighten’), alighten (‘alight, dismount’), amounten (‘amount, sum up’), aquikien (‘quicken, revive’), as[s]wagen (‘assuage, soothe’), atbresten (‘burst asunder, break open’), awilden (‘go wild, become wild’), bathen (‘bathe, wash’), blenden (‘blind’), blinden (‘blind’), blinnen (‘cease, stop’), boillen (‘boil’), brechen (‘break, violate’), breken (‘break’), brennen (‘burn’), bresten (‘burst, break’), chaungen (‘change’), cleven (‘cleave, split’), closen (‘close, shut’), craken (‘crack’), drenchen (‘drown, immerse’), drien (‘dry’), dronknen (‘drown, submerge’), droppen (‘drop’), drounen (‘drown’), encresen (‘increase’), fatten (‘fatten, enrich’), fillen (‘fill’), forbrennen (‘burn completely’), forfaren (‘perish, be destroyed’), frovren (‘freeze’), goden (‘improve, get better’), grouen (‘grow’), hardenen (‘harden’), helen (‘heal, cure’), hongen (‘hang’), ibeten (‘improve, make better’), ihelen (‘heal’), iturnen (‘turn, change’), iworthen (‘become, turn into’), kelen (‘cool, chill’), lenen (‘lend, grant; lean; make/become lean’), lighten (‘lighten, illuminate’), lightnen (‘lighten, illuminate’), litelen (‘belittle, make small’), louten (‘bow, bend’), louten (‘bow, bend’), melten (‘melt, dissolve’), multiplien (‘multiply’), openen (‘open, reveal’), quecchen (‘quench, extinguish’), quenchen (‘quench, extinguish’), ringen (‘ring, resound’), ripen (‘ripen, mature’), rosten (‘roast’), rungen (‘ring, resound’), sethen (‘boil, cook’), shiften (‘shift, change’), shitten (‘shut, close’), slaken (‘slacken, loosen’), sliden (‘slide, slip’), spillen (‘spill, destroy’), stillen (‘still, calm’), strecchen (‘stretch’), strengen (‘strengthen’), swellen (‘swell, grow larger’), swingen (‘swing, strike’), temen (‘empty,

void'), thestren ('thicken, darken'), tobreken ('break apart, shatter'), tobresten ('burst, break apart'), tocleven ('cleave, split apart'), toshiveren ('shiver, break into pieces'), unfolden ('unfold'), voiden ('void, empty'), wanen ('wane, diminish'), wansen ('decrease, diminish'), wersen ('worsen, deteriorate'), winden ('wind, twist'), wringen ('wring, twist'), writhen ('writhe, twist'),

- 30 motion verbs: areren ('rear, raise'), avauncen ('advance'), breden ('spread, breed'), departen ('depart, separate'), flitten ('move, shift'), gaderen ('gather'), heien ('hasten, make haste'), hevien ('lift, raise, make heavy'), hongen ('hang'), iwenden ('turn, return'), joinen ('join'), meven ('move, stir'), mouen ('move, shift'), reisen ('raise, rise'), reren ('raise, rear'), reren ('raise, rear'), risen ('rise'), samnen ('assemble, gather'), scateren ('scatter'), shaken ('shake, tremble'), sheten ('shoot'), sinken ('sink, descend'), sonderen ('separate, divide'), spreden ('spread'), staunchen ('staunch, stop the flow'), stoppen ('stop'), swenchen ('sway, swing'), turnen ('turn'), wenden ('turn, go'), withdrauen ('withdraw'),
- 18 Psych verbs: agasten ('frighten, terrify'), awaken ('awaken'), awakenen ('awaken'), forthinken ('regret, repent'), gladen ('gladden'), gladen ('gladden'), greven ('grieve, harm'), merveillen ('marvel, wonder'), peinen ('pain, torment'), pinen ('pine, suffer'), rejoisen ('rejoice'), repenten ('repent'), rosen ('rise, awaken'), waken ('wake, stay awake'), waken ('wake, stay awake'), wakenen ('awaken, become alert'), wratthen ('anger, enrage'), wrethen ('anger, enrage')
- 1 Aspectual verb: iswiken ('cease, stop')

C Materials for Section 4

C.1 Verbs for CHILDES analysis based on Middle English verbs

Cf. Section 4.3: ME CoS and motion verbs were translated to Present-Day English. The utterances containing these translations were extracted from CHILDES. Their "causative bias" was calculated based on occurrence of causative and anticausative structures in the CHILDES corpus. We list these 68 verbs with their causative bias:

- advance (causative bias = 0.00), awaken (0.20), bend (0.66), break (0.47), burn (0.53), burst (0.21), change (0.71), chill (0.08), clean (0.61), close (0.83), cool (0.39), crack (0.58), crush (0.60), darken (0.00), delight (0.50), diminish (0.00), dissolve (0.11), divide (0.63), drop (0.77), dry (0.55), ease (0.33), empty (0.88), enlarge (0.00), fade (0.08), fatten (0.40), fill (0.64), fold (0.77), freeze (0.27), grieve (0.00), grow (0.17), harden (0.29), hasten (0.50), heal (0.25), ignite (0.00), improve (0.22), increase (0.75), kindle (1.00), level (0.50), lighten (0.45), loosen (0.89), mature (0.25), melt (0.23), move (0.49), multiply (0.00), open (0.83), revolve (0.00), ripen (1.00), roll (0.57), shatter (0.14), shut (0.70), sink (0.18), slide (0.27), smooth (0.84), soften (0.80), spin (0.55), split (0.53), strengthen (1.00), stretch (0.47), submerge (0.00), swing

(0.26), tear (0.73), thicken (0.00), turn (0.67), twist (0.67), unfold (1.00), waken (0.40), weaken (1.00), wind (0.71).

References

Anderson, James M. 1973. *Structural aspects of language change* Longman Linguistics Library. London: Longman.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press. doi:10.2307/412595.

Austin, Alison C., Kathryn D. Schuler, Sarah Furlong & Elissa L. Newport. 2022. Learning a language from inconsistent input: Regularization in child and adult learners. *LLD* 18(3). doi:10.1080/15475441.2021.1954927. https://doi.org/10.1080/15475441.2021.1954927.

Baayen, Harald. 1993. On frequency, transparency and productivity. In *Yearbook of morphology 1992*, 181–208. Dordrecht: Springer. doi:10.1007/978-94-017-3710-4_7.

Bailey, Lyam M., Kate Lockary & Eve Higby. 2024. Cross-linguistic influence in the bilingual lexicon: Evidence for ubiquitous facilitation and context-dependent interference effects on lexical processing. *Bilingualism: Language and Cognition* 27(3). 495–514. doi:10.1017/S1366728923000597.

Belth, Caleb. 2024. A learning-based account of phonological tiers. *Linguistic Inquiry* 1–63.

Belth, Caleb, Sarah Payne, Deniz Beser, Jordan Kodner & Charles Yang. 2021. The Greedy and Recursive Search for Morphological Productivity. In Tecumseh Fitch, Claus Lamm, Helmut Leder & Kristin Teßmar-Raible (eds.), *Proceedings of the 43rd annual meeting of the cognitive science society*, 2869–2875. eScholarship.

Björnsdóttir, Sigríður Mjöll. 2021. Productivity and the acquisition of gender. *Journal of Child Language* 48(6). 1209–1234.

Bornstein, Marc H, Linda R Cote, Sharone Maital, Kathleen Painter, Sung-Yun Park, Liliana Pascual, Marie-Germaine Pêcheux, Josette Ruel, Paola Venuti & Andre Vyt. 2004. Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child development* 75(4). 1115–1139.

Bowerman, M. & William Croft. 2008. The Acquisition of the English Causative Alternation. In Melissa Bowerman & Penelope Brown (eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability*, 279–308. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bowerman, Melissa. 1974. Learning the structure of causative verbs: A study in the relationship of cognitive, semantic, and syntactic development. *Papers and Reports on Child Language Development* 8. 142–178.

Bowerman, Melissa. 1982. Evaluating explanations for children's semantic errors. In E. Wanner & L. R. Gleitman (eds.), *Language acquisition: The state of the art*, 296–323. Cambridge: Cambridge University Press.

Braginsky, Mika, Daniel Yurovsky, Virginia A Marchman & Michael C Frank. 2019. Consistency and variability in children's word learning across languages. *Open Mind* 3. 52–67.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins. doi:10.1075/tsl.9.

Chan, Erwin. 2008. *Structures and distributions in morphological learning*. Philadelphia, PA: University of Pennsylvania dissertation.

Chu, Junyi, Pierina Cheung, Rose M Schneider, Jessica Sullivan & David Barner. 2020. Counting to infinity: Does learning the syntax of the count list predict knowledge that numbers are infinite? *Cognitive Science* 44(8). e12875.

Comrie, Bernard. 2006. Transitivity pairs, markedness, and diachronic stability. *Linguistics* 44(2). 303–318.

Cournane, Ailís. 2017. In defense of the child innovator. In Eric Mathieu & Robert Truswell (eds.), *Micro change and macro change in diachronic syntax*, 10–24. Oxford: Oxford University Press. doi:10.1093/oso/9780198747840.003.0002.

Demirdache, Hamida & Fabienne Martin. 2015. Agent control over non culminating events. In Elisa Barrajón López, José Luis Cifuentes Honrubia & Susana Rodríguez Rosique (eds.), *Verb classes and aspect* 9, 185–217. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Dietz, Klaus. 2004. Die altenglischen Präfixbildungen und ihre Charakteristik. *Anglia* (122(4)). 561–613.

Dolatian, Hossep & Jordan Kodner. in press. Analogical Extension from Irregular Paradigms in Iranian Armenian: A Case of "Elsewhere Reversal". *Diachronica* in press.

Durkin, Philip. 2014. *Borrowed words. A history of loan words in english*. Oxford: Oxford University Press.

Elenbaas, Marion. 2007. *The synchronic and diachronic syntax of the English verb-particle combination* (LOT 149). Utrecht: LOT.

Elman, Jeffrey. 1998. Generalization, simple recurrent networks, and the emergence of structure. In *Proceedings of the twentieth annual conference of the cognitive science society*, 543–548. Mahwah, NJ: Lawrence Erlbaum.

Emond, Emeryse & Rushen Shi. 2020. Infants' rule generalization is governed by the Tolerance Principle. In *45th annual boston university conference on language development*, .

Fenson, Larry, Philip S Dale, J Steven Reznick, Elizabeth Bates, Donna J Thal, Stephen J Pethick, Michael Tomasello, Carolyn B Mervis & Joan Stiles. 1994. Variability in early communicative development. *Monographs of the society for research in child development* 59(5). 1–185.

Fotiadou, Georgia & Ianthi Maria Tsimpli. 2010. The acquisition of transitivity alternations in Greek: Does frequency count? *Lingua* 120(11). 2605–2626. doi:10.1016/j.lingua.2010.06.011.

García García, Luisa. 2019. The basic valency orientation of Old English and the causative *ja-* formation: A synchronic and diachronic approach. *English Language and Linguistics* 24(1). 153–177. doi:10.1017/S1360674318000345.

García García, Luisa & Richard Ingham. 2023. Language Contact Effects on Verb Semantic Classes: Lability in Early English and Old French. In Sara M. Pons-Sanz & Louise Sylvester (eds.), *Medieval English in a Multilingual Context*, 343–375. Cham: Springer International Publishing. doi:10.1007/978-3-031-30947-2_12.

van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the language faculty*.

Oxford University Press.
van Gelderen, Elly. 2018. *The diachrony of verb meaning. Aspect and argument structure*. Routledge.
Gleitman, Lila. 1990. The structural sources of verb meanings. *Language Acquisition* 1. 3–55.
Goodman, Judith C, Philip S Dale & Ping Li. 2008. Does frequency count? Parental input and the acquisition of vocabulary. *Journal of child language* 35(3). 515–531.
Guilbeault, Douglas, Spencer Caplan & Charles Yang. 2026. A simple threshold captures the social learning of conventions. *Proceedings of the National Academy of Sciences* 123(17). e2508061123. doi:10.1073/pnas.2508061123. https://www.pnas.org/doi/abs/10.1073/pnas.2508061123.
Hart, Betty & Todd R Risley. 1995. *Meaningful differences in the everyday experience of young american children*. Towson, MD: Paul H Brookes Publishing.
Hart, Betty & Todd R Risley. 2003. The early catastrophe: The 30 million word gap by age 3. *American Educator* 27(1). 4–9.
Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity* (Studies in Language Companion Series 23), 87–120. Amsterdam: J. Benjamins.
Heidinger, Steffen. 2010. *French anticausatives: A diachronic perspective*. Berlin: de Gruyter.
Heine, Bernd & Tania Kuteva (eds.). 2005. *Language contact and grammatical change* Cambridge Approaches to Language Contact. Cambridge: Cambridge Univ. Press.
Henke, Ryan E. 2023. Rules and exceptions: A tolerance principle account of the possessive suffix in northern east cree. *Journal of Child Language* 50(5). 1119–1154.
Hudson Kam, Carla L & Elissa L Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1(2). 151–195.
Ingham, Richard. 2012. *The transmission of Anglo-Norman: Language history and language acquisition*, vol. 9 Language Faculty and Beyond. Amsterdam, Philadelphia: John Benjamins.
Ingham, Richard. 2020. How Contact with French Drove Patient-Lability in English. *Transactions of the Philological Society* 118(3). 447–467.
Inglese, Guglielmo. 2025. Diachronic pathways to reflexive-only verbs in Italian. *Italian Journal of Linguistics* 17. to appear.
Irani, Ava. 2019. *Learning from positive evidence: The acquisition of verb argument structure*. Philadelphia: University of Pennsylvania dissertation.
Jelinek, Frederick. 1997. *Statistical methods for speech recognition*. MIT press.
Kanampiu, Patrick Njue, Alexander Martin & Jennifer Culbertson. 2025. Semantic and morphophonological productivity in the kîîtharaka gender system: A quantitative study. *Glossa: a journal of general linguistics* 10(1).
Kern, Sophie. 2003. Le compte-rendu parental au service de l'évaluation de la production lexicale des enfants français entre 16 et 30 mois. *Glossa* (85). 48–61.
Kibbee, Douglas A. 1991. *For to Speke French Trewely: The French Language in England, 1000–1600: Its Status, Description and Instruction*. Amsterdam: Benjamins.
Kiparsky, Paul. 1965. *Phonological change*. Cambridge, MA: Massachusetts Institute

of Technology dissertation.
Kiparsky, Paul. 1973. Elsewhere in phonology. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for morris halle*, 93–106. New York: Holt, Rinehart and Winston.
Kodner, Jordan. 2019a. Estimating child linguistic experience from historical corpora. *Glossa-an International Journal of Linguistics* 4 (1)(122). 1–14.
Kodner, Jordan. 2019b. Estimating child linguistic experience from historical corpora. *Glossa: a journal of general linguistics* 4(1).
Kodner, Jordan. 2020. *Language acquisition in the past*. Philadelphia: University of Pennsylvania dissertation.
Kodner, Jordan. 2023a. *Laissez-Faire* Analogical Change. *Journal of Historical Syntax* 7. 6–19.
Kodner, Jordan. 2023b. What learning Latin verbal morphology tells us about morphological theory. *Natural Language & Linguistic Theory* 41(2). 733–792. doi:10.1007/s11049-022-09553-2.
Kodner, Jordan, Sarah Payne, Salam Khalifa & Zoey Liu. 2025. Evaluating Learning Trajectories of Neural Morphology Acquisition Models. *Linguistics Vanguard* doi:10.1515/lingvan-2024-0215.
Koulaguina, Elena & Rushen Shi. 2019. Rule generalization from inconsistent input in early infancy. *Language Acquisition* 26(4). 416–435.
Kroch, Anthony. 2005. Modeling language change and language acquisition. In *Expansion of an lsa institute forum lecture*, .
Kroch, Anthony & Ann Taylor (eds.). 2000. *The Penn-Helsinki Parsed Corpus of Middle English, Second edition* (*PPCME2*). Philadelphia: University of Pennsylvania.
Labov, William. 1972. Some principles of linguistic methodology. *Language in society* 1(1). 97–120. doi:10.1017/s0047404500006576.
Labov, William. 2007. Transmission and diffusion. *Language* 83(2). 344–387.
Labov, William, Malcah Yaeger & Richard Steiner. 1972. *A quantitative study of sound change in progress*, vol. 1. US Regional Survey.
Letuchiy, Alexander. 2010. Lability and Spontaneity. In Patrick Brandt & Marco García García (eds.), *Transitivity: Form, Meaning, Acquisition, and Processing*, vol. 166 Linguistik Aktuell/Linguistics Today, 237–255. Amsterdam: John Benjamins Publishing Company. doi:10.1075/la.166.
Levin, Beth. 1993. *English Verb Classes and Alternations. A Preliminary Investigation*. Chicago: University of Chicago Press.
Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface* (Linguistic Inquiry Monographs 26). Cambridge, Mass: MIT Press.
Li, Daoxin. 2024. *Distributional learning of syntactic generalizations*. Philadelphia: University of Pennsylvania Doctoral thesis.
Li, Daoxin & Kathryn D. Schuler. 2023. Acquiring recursive structures through distributional learning. *Language Acquisition* 30(3-4). 323–336. doi:10.1080/10489223.2023.2185522.
Lightfoot, David W. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.
MacWhinney, Brian. 1978. The acquisition of morphophonology. *Monographs of the*

Society for Research in Child Development 1–123. doi:10.2307/1166047.
MacWhinney, Brian. 2000. *The childes project. 1: Transcription format and programs*. Mahwah: Erlbaum.
McMillion, Alan. 2006. *Labile Verbs in English: Their Meaning, Behavior and Structure*. PhD Thesis Stockholm: University of Stockholm.
Müller, Natascha (ed.). 2003. *(in)vulnerable domains in multilingualism*. Amsterdam: John Benjamins.
Nagy, William E & Richard C Anderson. 1984. How many words are there in printed school English? *Reading Research Quarterly* 304–330. doi:10.2307/747823.
Nardy, Aurélie, Jean-Pierre Chevrot & Stéphanie Barbu. 2014. Sociolinguistic convergence and social interactions within a group of preschoolers: A longitudinal study. *Language Variation and Change* 26(3). 273–301.
Niyogi, Partha & Robert C Berwick. 1997. A dynamical systems model for language change. *Complex Systems* 11(3). 161–204.
Nowenstein, Iris, Sigríður Sigurjónsdóttir, Charles Yang, Anton Ingason & Joel Wallenberg. 2020. The meaning of case: Morphosyntactic bootstrapping and icelandic datives. In Megan M. Brown & Alexandra Kohut (eds.), *45th Annual Boston University Conference on Language Development*, 402–415. Sommerville, MA: Cascadilla Press.
Palermo, David S & V Lynn Eberhart. 1968. On the learning of morphological rules: An experimental analogy. *Journal of Verbal Learning and Verbal Behavior* 7(2). 337–344.
Paul, Hermann. 1880. *Prinzipien der Sprachgeschichte*. T ubingen. Niemeyer.
Payne, Sarah & Charles Yang. 2023. Making good on bads: Commentary on dressler et al. 2020. *Italian Journal of Linguistics* .
Pearl, Lisa & Jon Sprouse. 2018. The acquisition of linking theories: A tolerance principle approach to learning utah and rutah. *Irvine: University of California, Irvine* .
Percillier, Michael, Yela Schauwecker, Achim Stein & Carola Trips. 2024. *Carrying Verbs Across the Channel - Modelling Change in Bilingual Medieval England*. Palgrave Macmillan.
Pierrehumbert, Janet B. 2003. On frequency, transparency and productivity. In *Probabilistic linguistics*, 177–228. Cambridge, MA: MIT Press.
Pinker, Steven. 1989. *Learnability and Cognition. The Acquisition of Argument Structure*. Cambrige, MA: MIT Press.
Pinker, Steven. 2013. *Learnability and Cognition. The Acquisition of Argument Structure*. Cambrige, MA: MIT Press 2nd edn.
Richter, Caitlin L. 2021. *Alternation-sensitive phone learning: Implications for children's development and language change*: University of Pennsylvania dissertation.
Ringe, Don & Ann Taylor. ???? *The development of old english*. Oxford University PressOxford 1st edn. doi:10.1093/acprof:oso/9780199207848.001.0001. https://academic.oup.com/book/36266.
Rothwell, William. 1980. Lexical borrowing in a medieval context. *Bulletin ofthe John Rylands Library* 63. 118–43.
Schäfer, Florian. 2008. *The Syntax of (Anti-)Causatives. External arguments in change-of-state contexts*. Amsterdam, Philadelphia: Benjamins.

Schuler, Kathryn D. 2017. *The acquisition of productive rules in child and adult language learners*: Georgetown University dissertation.

Schuler, Kathryn D, Charles Yang & Elissa Newport. 2016. Testing the tolerance principle: Children form productive rules when it is more computationally efficient to do so. In *The 38th cognitive society annual meeting*, 2321–26. Mahwah, NJ: Lawrence Erlbaum.

Scott, Rose M. & Cynthia Fisher. 2009. Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes* 24(6). 777–803. doi:10.1080/01690960802573236.

Shi, Rushen & Emeryse Emond. 2023. The threshold of rule productivity in infants. *Frontiers in Psychology* 14. 1251124. doi:10.3389/fpsyg.2023.1251124.

Short, Ian. 1995. Tam Angli quam Franci: Self-Definition in Anglo-Norman England. *Anglo-Norman Studies* 153–175.

Short, Ian. 2007. *Manual of Anglo-Norman*. London: Anglo-Norman Text Society.

Stein, Achim. 2026. From CHAT to Coded CoNLL-U: A Reproducible Pipeline for the Syntactic Annotation and Querying of Child Language Data. In Nicoletta Calzolari et al. (eds.), *Proceedings of the 15th Language Resources and Evaluation Conference (LREC), May 11-16, 2026*, European Language Resources Association (ELRA). doi:10.63317/498zo5heasd5.

Straka, Milan. 2018. UDPipe 2.0 prototype at CoNLL 2018 UD shared task. In *Proceedings of the CoNLL 2018 shared task: Multilingual parsing from raw text to universal dependencies*, 197–207. Brussels, Belgium: Association for Computational Linguistics. doi:10.18653/v1/K18-2020.

Swingley, Daniel & Colman Humphrey. 2018. Quantitative linguistic predictors of infants' learning of specific English words. *Child development* 89(4). 1247–1267.

Szagun, Gisela, Claudia Steinbrink, Melanie Franik & Barbara Stumper. 2006. Development of vocabulary and grammar in young German-speaking children assessed with a German language development inventory. *First language* 26(3). 259–280.

Trips, Carola. 2014. The position proper of the adjective in Middle English: A result of language contact? In Harry Perridon, Petra Sleeman & Freek Van de Velde (eds.), *Adjectives in germanic and romance*, 73–94. Amsterdam/Philadelphia: Benjamins Publishing.

Trips, Carola. 2020. Copying of argument structure: A gap in borrowing scales and a new approach to model contact-induced change. In Bridget Drinka (ed.), *Historical linguistics 2017. Selected papers from the 23rd international conference on historical linguistics, san antonio, texas, 31 july - 4 august 2017* (CILT 350), 413–434. Benjamins Publishing.

Trips, Carola & Tom Rainsford. 2022. Tolerating subject-Experiencers? Yang's Tolerance Principle applied to psych verbs under contact in Middle English. *Journal of Historical Syntax* 6(12). 1–43.

Trips, Carola, Tara Struik & Lena Kaltenbach. 2026. Language Contact and Language Change: Contact-Induced Structural Change in Medieval England. In *Reference Module in Social Sciences*, B9780323955041008115. Elsevier. doi:10.1016/B978-0-323-95504-1.00811-5.

Truswell, Robert et al. (eds.). 2018. *A Parsed Linguistic Atlas of Early Middle English*.

Edinburgh: Angust McIntosh Centre for Historical Linguistics.
van Hout, Angeliek. 2017. Agent control and the acquisition of event culmination in Basque, Dutch, English, Spanish and Mandarin". In Maria LaMendola & Jennifer Scott (eds.), *BUCLD 41: Proceedings of the 41st annual Boston University Conference on Language Development*, 323–332. Somerville, MA: Cascadilla Press.
Visser, Fredericus Theodorus. 1963. *An historical syntax of the english language*. E. J. Brill.
Walkden, George. 2019. The many faces of uniformitarianism in linguistics. *Glossa: a journal of general linguistics* 4(1). 52. doi:10.5334/gjgl.888.
Yang, Charles. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.
Yang, Charles. 2005. On productivity. *Linguistic Variation Yearbook* 5. 265–302. doi:10.1075/livy.5.09yan.
Yang, Charles. 2013. Who's afraid of George Kingsley Zipf? Or: Do children and chimps have language? *Significance* 10(6). 29–34. doi:10.1111/j.1740-9713.2013.00708.x.
Yang, Charles. 2016. *The price of linguistic productivity*. Cambridge, Mass.: MIT Press.
Yang, Charles. 2018. A user's guide to the tolerance principle. LingBuzz preprint. http://ling.auf.net/lingbuzz/004146.
Yang, Charles. 2024. Phonological regularity and breakdown. an account of vowel length leveling in middle english. In *The method works: Studies on language change in honor of don ringe*, 237–259. Springer.
Yang, Charles & Silvina Montrul. 2017. Learning datives: The tolerance principle in monolingual and bilingual acquisition. *Second Language Research* 33(1). 119–144.
Zimmermann, Richard (ed.). 2014. *Parsed corpus of Middle English poetry.* Geneva.
Zipf, George Kingsley. 1936. *The psycho-biology of language: An introduction to dynamic philology*. Milton Park, Abingdon, UK: Routledge.

Jordan Kodner
Department of Linguistics
Stony Brook University
Stony Brook, NY 11794-4376, USA
jordan.kodner@stonybrook.edu
orcid.org/0000-0001-8632-9859

Achim Stein
Institut für Linguistik/Romanistik
Universität Stuttgart
Keplerstrasse 17
70174 Stuttgart, Germany
achim.stein@ling.uni-stuttgart.de
orcid.org/0000-0001-8848-1977

Carola Trips
Anglistik IV
Universität Mannheim
B6, 30-32
68159 Mannheim, Germany
carola.trips@uni-mannheim.de
orcid.org/0009-0001-8224-3107