

New methods for old languages: the comparability of data

W13, ICHL26, Heidelberg

Alessia Cassarà (Universität Stuttgart); Lena Kaltenbach (Universität Mannheim);
Mariapaola Piccione (Universität Stuttgart); Tara Struik (Universität Mannheim)

Corresponding convenor: tara.struik@uni-mannheim.de

While historical linguistics is traditionally known to suffer from a “bad data problem” (Labov 1994: 11), the field has seen a surge in the development of (annotated) data collections and computational tools to trace quantitative changes throughout the history of languages, allowing researchers to get more out of the (often sparse) data than ever before. This availability of data opens up many new avenues for research, in particular in explaining the cognitive mechanisms behind language change. In this workshop we want to bring together researchers working in different disciplines to discuss novel empirical methods that allow us to investigate the relation between the structural changes we observe in historical texts and the factors which arguably led to these changes. We aim to do this by focusing on a) how historical corpus data can be related to **models of language learning**, b) **contemporary psycholinguistic models** and c) how we can deal with the **heterogeneity of historical data** in relation to these models.

Historical linguists have discussed the link between historical change and changes in the input and have proposed models that make use of psycholinguistic explanations, especially in terms of language acquisition (e.g. Lightfoot 1999, 2017). However, a challenge for the study of the role of language acquisition in language change is that there is no direct access to the input for past stages of languages. Approximating the input by using corpora of child-directed speech (e.g. CHILDES) for contemporary languages has resulted in the development of learning models, which may also be informative for the historical stages. For instance, Yang’s (2016) Tolerance Principle has been shown to work effectively with small amounts of data, making it very attractive for historical work (Kodner 2020, 2022; Drescher and Lahiri 2022, Ringe and Yang 2022, Trips and Rainsford 2022). However, the application of such models on corpus data requires careful consideration of how the data obtained from corpora can be compared to the input a child received (cf. Trips and Rainsford 2022 for discussion). One potential solution is to compare the frequencies of the most common verbs in a corpus to the most common verbs in a sample of child-directed speech, as Kodner (2019) demonstrates that there is a substantial overlap.

From both a psycholinguistic and historical linguistic perspective the relationship between language change and mechanisms of language processing has only rarely been explicitly addressed (for exceptions, see Jäger & Rosenbach 2008; de Smet & de Velde 2017; see also the contributions in Hundt et al. 2017 and the ongoing work by the the DFG Research Unit SILPAC (FOR 5157)). Notably, some authors have recently pointed to the importance of cross-linguistic and within-language structural priming and syntactic adaptation for studies

of (contact-induced) language change (e.g. Pickering & Garrod 2017; Kaan & Chun 2018; Kootstra & Şahin 2018; Kootstra & Muysken 2019). Effects of priming may be observable in historical corpora in the form of persistence of linguistic forms (see Ecay and Tamminga 2017; also Gries 2005; Szmrecsanyi 2006). From a Uniformitarian perspective (see Bergs 2012, Walkden 2019 for discussion), it follows that psycholinguistic processes active in language change should not differ fundamentally across languages or language stages. Methodologically, changes observed in diachrony could in principle also be elicited in psycholinguistic experiments and the results and methods of psycholinguistic experiments could inform historical corpus analyses.

Applying psycholinguistic methods and learning models to historical data also requires us to think critically about the nature of our data and how informative they are about the actual linguistic environment in which language acquisition and change takes place. Historical corpora may be heterogeneous in nature, consisting of many different genres (e.g. legal prose, narrative verse, etc.), which may not all be equally representative of a language user's input. Some types of text, e.g. theatrical texts, conversation manuals, direct speech in verse narratives, etc. have been argued to be particularly close to spoken language in the past (Ernst 1980, Ayres-Bennett 2000); also, it has been shown that language change does not proceed at the same rate in all text genres (Whitt 2018). However, it is not clear whether a restrictive approach to selecting corpus texts is preferable to one which instead draws on as much data as possible, using statistical techniques to evaluate the effect of genre. A further open question is the extent to which the writers of historical texts are themselves influenced by mechanisms such as priming, whether it is self-priming within a single text, between the two writers in private correspondence or even between two languages in translations. Similarly, it is not always clear what the impact of the linguistic background of individual authors is on the output – are they, for instance, monolinguals, early bilinguals, or possibly late bilinguals writing in their first language or late bilinguals writing in their second language?

In this workshop, we aim to compare different types of historical corpus data not only with each other, but also with the input to language acquirers and with data elicited in psycholinguistic experiments in order to develop novel methodologies bringing the fields of historical linguistics, psycholinguistics and language acquisition closer together. We invite contributions which answer or relate to the following research questions and topics:

- How can models of learnability be applied to historical data?
- What are the psycholinguistic processes behind historical language change?
- Which insights does historical linguistics provide for the study of these psycholinguistic processes?
- Which methods and resources are the best to use if we want to relate historical data to language learner input and which are best for researching the relationship between experimental data and historical data?
- Which additional data types/methodologies can contribute to bridging the gap between the disciplines of historical linguistics, acquisition studies and psycholinguistics, e.g. artificial language studies, longitudinal studies, computational models of language change, etc.?

- How can insights from historical sociolinguistics and philology contribute to a better understanding of the heterogeneity of historical corpus data and the linguistic background of individual authors?
- To what extent are the writers of historical texts themselves influenced by mechanisms of language processing, such as intra- and interindividual priming in monolingual and bilingual situations? How can we use notions such as persistence in historical corpora to tap into the cognitive processes behind the text production of medieval authors?

References

- Ayres-Bennett, W. 2000. Voices from the past: Sources of seventeenth-century spoken French. *Romanische Forschungen*, 112, 323–348.
- Bergs, A. 2012. The Uniformitarian Principle and the risk of anachronisms in language and social history. In J. M. Hernández-Campoy & J. C. Conde-Silvestre (Eds.), *The handbook of historical sociolinguistics* (1st ed., pp. 83–101). <https://doi.org/10.1002/9781118257227.ch5>
- De Smet, H., & Van de Velde, F. 2017. Experimenting on the past: A case study on changing analysability in English ly-adverbs. *English Language & Linguistics*, 21(2), 317–340. <https://doi.org/10.1017/S1360674317000168>
- Dresher, B. E., & Lahiri, A. 2022. The foot in the history of English: Challenges to metrical coherence. In B. Los, C. Cowie, P. Honeybone, & G. Trousdale (Eds.), *English historical linguistics: Change in structure and meaning: Papers from the XXth ICEHL* (pp. 42–59). John Benjamins Publishing Company.
- Ecay, A., & Tamminga, M. 2017. Persistence as a diagnostic of grammatical status: The case of Middle English negation. In E. Mathieu & Truswell, B. (Eds.). *Micro-change and Macro-change in diachronic syntax* (pp. 202–215). Oxford University Press. <https://doi.org/10.1093/oso/9780198747840.003.0013>
- Ernst, G. 1980. Prolegomena zu einer Geschichte des gesprochenen Französisch. In H. Stimm (Ed.), *Zur Geschichte des gesprochenen Französisch und zur Sprachlenkung im Gegenwartsfranzösischen* (pp. 1–14). Steiner.
- Gries, S. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research*, 34(4), 365–399.
- Hundt, M., Mollin, S., & Pfenninger, S. (Eds.). 2017. *The changing English language: Psycholinguistic perspectives*. Cambridge University Press.
- Jäger, G., & Rosenbach, A. 2008. Priming as a testing ground for historical linguists? - A reply to Chang, Eckardt, and Traugott. *Theoretical Linguistics*, 34(2), 143–156.
- Kaan, E., & Chun, E. 2018. Priming and adaptation in native speakers and second-language learners. *Bilingualism: Language and Cognition*, 21(2), 228–242. <https://doi.org/10.1017/S1366728916001231>
- Kodner, J. 2019. Estimating child linguistic experience from historical corpora. *Glossa. A Journal of General Linguistics*, 4(1), 122.
- Kodner, J. 2020. *Language acquisition in the past* [Doctoral dissertation]. University of Pennsylvania.
- Kodner, J. (2022). What learning Latin verbal morphology tells us about morphological theory. *Natural Language and Linguistic Theory*. <https://doi.org/10.1007/s11049-022-09553-2>

- Kootstra, G. J., & Muysken, P. 2019. Structural priming, levels of awareness, and agency in contact-induced language change. *Languages*, 4(3), 65. <https://doi.org/10.3390/languages4030065>
- Kootstra, G. J., & Şahin, H. 2018. Crosslinguistic structural priming as a mechanism of contact-induced language change: Evidence from Papiamentto-Dutch bilinguals in Aruba and the Netherlands. *Language*, 94(4), 902–930.
- Labov, W. 1994. *Principles of linguistic change, Volume 1: Internal factors*. Blackwell.
- Lightfoot, D. 1999. *The development of language: Acquisition, change, and evolution*. Blackwell.
- Lightfoot, D. 2017. Acquisition and learnability. In A. Ledgeway & I. Roberts (Eds.), *The Cambridge handbook of historical syntax* (pp. 381–400). Cambridge University Press.
- Pickering, M., & Garrod, S. 2017. Priming and language change. In M. Hundt, S. Mollin, & S. Pfenninger (Eds.), *The changing English language: Psycholinguistic perspectives* (pp. 173–190). Cambridge University Press.
- Ringe, D., & Yang, C. 2022. The threshold of productivity and the ‘irregularization’ of verbs in Early Modern English. In B. Los, C. Cowie, P. Honeybone, & G. Trousdale (Eds.), *English historical linguistics: Change in structure and meaning: Papers from the XXth ICEHL* (pp. 91–112). John Benjamins Publishing Company.
- Szmrecsanyi, B. 2006. *Morphosyntactic persistence in spoken English: A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. Mouton de Gruyter.
- Trips, C., & Rainsford, T. 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.
- Walkden, G. 2019. The many faces of uniformitarianism in linguistics. *Glossa: A Journal of General Linguistics*, 4(1), 1. <https://doi.org/10.5334/gjgl.888>
- Whitt, R. J. (Ed.). 2018. *Diachronic corpora, genre, and language change*. John Benjamins Publishing Company.
- Yang, C. 2016. *The price of linguistic productivity: How children learn to break the rules of language*. MIT Press.

Dative Experiencer Psych Verbs in (Old) French

Lisa Figura, lisa.figura@uni-hamburg.de
Universität Hamburg

Psych Verbs (PV) are verbs that express mental or emotional concepts and have an argument bearing the theta role Experiencer. PV represent a phenomenon that has been discussed for decades (cf. Hirsch 2018 for an overview) as they are a heterogeneous verb class with an unstable argument structure and different syntactic constructions. There are large differences within the (Romance) languages, synchronically as well as diachronically. According to Belletti/Rizzi (1988), the Dative Experiencer Psych Verbs (Dat.Exp.PV) form one of three classes of PV. This class is characterised by the fact that the dative objects can occupy a preverbal position. This non-canonical use is possible in Latin and Romance Languages, e.g. in Spanish, as in (1):

- (1) **A Ana** le **gusta** el chocolate.
 DAT.EXP CL.DAT like.3SG the chocolate
 “Anna likes chocolate.”

In Old French, the preverbal Dat.Exp can also be found (cf. Mathieu 2006: 2), as in (2):

- (2) Et bien set qu’ **a sa mere plect** que [...]
 And well know.3SG that DAT.EXP like.3SG
 “And she knows well that it is her mother's will that [...]”

In Modern French, this structure is ungrammatical: the Dat.Exp can occupy the preverbal position only by topicalisation or dislocation with doubling (cf. Fischer 2019), as in (3):

- (3) **A Marie,** la musique classique lui **plaît.**
 DAT.EXP the music classic PRN.DAT like.3SG
 “Marie likes classic music.”

In Old French, however, the non-canonical use of Dat.Exp.PV is common (cf. Mathieu 2006), which raises the question why it has disappeared.

The talk opens a new perspective in the debate on Dat.Exp.PV linking language change to principles of language acquisition. My hypothesis is twofold, assuming two parallel developments of Dat.Exp.PV that together lead to its gradual loss – with a few exceptions where the Experiencer is grammaticalised in object position. The first part of my hypothesis is based on markedness in terms of structural complexity: it is assumed that the structure of Dat.Exp.PV is more complex than non-PV verb classes and not acquired easily. The second part of my hypothesis is based on computational efficiency and the assumption that during first language acquisition, rules and exceptions are organised to optimise linguistic

processing. Assuming the Tolerance Principle (cf. Yang 2016), it is argued that preverbal Dat.Exp have not been acquired as a productive rule due to the amount of exceptions to this rule.

On the one hand, it is argued that the Old French PV did not undergo the expected developments towards intransitivity and stativity, which are generally considered to be the properties of less marked argument and event structures (cf. Van Gelderen 2014, 2019; Batllori et al. 2019). Preverbal Dat.Exp can be regarded as a marked input because of the irregular theta-role mapping (cf. Scontras et al. 2015). Language acquisition research shows that marked options are acquired later (cf. Roberts 2007, among others) and Schmitz (2006) argues that dative case is more difficult to be acquired than other cases. On the other hand, I will show that the fixation of French word order from OV to VO during the 12th century led to a low frequency of Dat.Exp in subject position. I will argue that this is the reason why Dat.Exp grammaticalised in object position. This process will be explained by referring to the Tolerance Principle, which has already been applied to Middle English PV and their argument structure (cf. Trips/Rainsford 2022). I will suggest that Old French language learners could not maintain a productive rule which provided a preverbal and a postverbal position for Dat.Exp. Since preverbal Dat.Exp were not as frequent in the PLD as postverbal Dat.Exp, the learners hypothesised as the productive rule for Dat.Exp only the postverbal position.

These hypotheses will be tested by examining two Old French corpora (*MCVF-PPCHF* and the *Nouveau Corpus d'Amsterdam*). Frequencies of both preverbal and postverbal Dat.Exp will be gathered and their argument and event structure will be analysed. A first pretest analysing the PV *plaire* in the *MCVF-PPCHF* showed 268 occurrences of this verb with a Dat.Exp, of which 115 are preverbal and 153 are postverbal. Further results – also concerning the event and argument structure of different Dat.Exp – will be presented in the talk.

References

- Batllori, Montserrat / Gibert-Sotelo, Elisabeth / Pujol, Isabel (2019): “Changes in the argument and event structure of psych verbs in the history of Spanish”, in Bouzouita, Miriam / Breitbarth, Anne /
- Belletti, Adriana / Rizzi, Luigi (1988): „Psych-Verbs and θ -Theory“. *Natural Language & Linguistic Theory*, 6(3), 291–352.
- Fischer, Susann (2019): „On oblique experiencers in psych verb constructions“, in Pomino, Natascha (ed.): *Proceedings of the VII Nereus International Workshop “Morphosyntactic and semantic aspects of the DP in Romance and beyond”*. Arbeitspapier 131. Fachbereich Linguistik, Universität Konstanz 2019, 109-128.
- Hirsch, Nils (2018): *German Psych-Verbs: Insights from a Decompositional Perspective*. Dissertation, Humboldt-Universität zu Berlin, Berlin.

- Kroch, Anthony / Santorini, Beatrice (2021): Penn-BFM Parsed Corpus of Historical French, version 1.0. <https://github.com/beatrice57/mcvf-plus-ppchf>
- Martineau, France, Paul Hirschbühler, Anthony Kroch, and Yves Charles Morin (2021): MCVF Corpus, parsed, version 2.0. <https://github.com/beatrice57/mcvf-plus-ppchf>
- Mathieu, Eric (2006): “Quirky Subjects in Old French“. *Studia Linguistica* 60(3): 282 – 312.
- Roberts, Ian (2007): *Diachronic Syntax*. Oxford: Oxford University Press.
- Schmitz, Katrin (2006): „Indirect objects and dative case in monolingual German and bilingual German/Romance language acquisition“, in Hole, Daniel / Meinunger, Andre / Abraham, Werner (eds.): *Datives and Other Cases*. Amsterdam/Philadelphia: John Benjamins., 239-268.
- Scontras, Gregory / Fuchs, Zuzanna / Polinsky, Maria (2015): „Heritage language and linguistic theory“, *Frontiers in Psychology*. 6:1545.
- Stein, Achim et al. (2006). *Nouveau Corpus d'Amsterdam. Corpus informatique de textes littéraires d'ancien français (ca. 1150-1350)*, établi par Anthonij Dees (Amsterdam 1987), remanié par Achim Stein, Pierre Kunstmann et Martin-D. Gleßgen. Stuttgart: Institut für Linguistik/Romanistik, version 3.
- Trips, Carola / Rainsford, Thomas (2022): “Tolerating subject-experiencers? Yang’s Tolerance Principle applied to psych verbs under contact in Middle English”, *Journal of Historical Syntax*, Vol. 6, No. 12. 1-43.
- Van Gelderen, Elly (2019): “Stability and Change in Intransitive Argument Structure“, *Open Linguistics*, vol. 5, no. 1: 217-232. Wolfsgruber, Anne C. (2017): “On, Se and related valency alternations in Medieval French”, *Lingvisticae Investigationes*, 40: 59-80.
- Van Gelderen, Elly (2014): “Changes in Psych-verbs: A reanalysis of little v.”, *Journal of Catalan Linguistics*, 13: 99-122.
- Yang, Charles D. (2016): *The price of linguistic productivity*. MIT Press.

How to use Yang's Principles to model acquisition in diachrony The case of psych verbs

Carola Trips¹ & Tom Rainsford²

¹University of Mannheim, ²University of Stuttgart

Any study looking at acquisition in the past must infer aspects of the input to which children are exposed from written texts, yet it is clear that they are not equivalent. As a model of learnability, Yang's (2016) Tolerance and Sufficiency Principles are good candidates for the study of the acquisition of productive rules in historical data, and they have been applied in a number of recent studies (e.g. Kodner 2019, 2020, 2022; Drescher and Lahiri 2022, Ringe and Yang 2022). The model provides a simple but effective algorithm for predicting the point at which language learners will conclude that the number of lexical items belonging to a given class and providing positive evidence for a particular rule is sufficient to conclude that the rule is productive, barring a small number of exceptions which must be memorized. On the basis of child-directed speech data from the CHILDES corpus (MacWhinney 2000), Yang (2016) has shown that the Tolerance Principle is effective in modelling the acquisition of productive rules such as the use of the *-ed* past-tense marker and its corollary, the Sufficiency Principle, is well-suited to the acquisition of argument structure, such as modelling the subset of ditransitive verbs showing double object constructions in modern English (see also Kodner 2019). Not only has the Sufficiency Principle been shown to be effective in correctly predicting the course of acquisition from small amounts of data, similar to those to which a child would be exposed and to the limited data available to historical linguists, the calculation only requires two parameters to be estimated: the total number of lexical items within the class to which the learner is exposed (henceforth N) and the number of these lexical items to which the rule in question can be applied (henceforth M).

However, applying the Sufficiency Principle to historical data brings a number of unique problems not present in the child-directed speech data examined by Yang. In a recent study of the acquisition of psych verbs in Middle English, Trips and Rainsford (2022) identify three central issues: First is the class size problem: how is it possible to estimate the number of lexical items in a particular class (N), in this case psych verbs, from heterogeneous historical corpora? Second is the attestation problem: what is the best way to estimate the positive evidence for a given rule (M), in this case, the use of a subject-EXPERIENCER argument, from historical data? Third is the data compatibility problem: to what extent is data from historical texts comparable to child-directed speech data?

In the present article, we re-examine the validity of the assumptions made by the authors to address these problems. First, contrasting the psych verbs attested in sections M3 (1350-1420) and M4 (1420-1500) of the *Penn-Helsinki Parsed Corpus of Middle English* (PPCME2) and those attested in modern English child-directed speech from the CHILDES corpus, we show that there is broad semantic equivalence between the most frequent verbs in historical texts and those found in child-directed speech, confirming that basing Sufficiency Principle calculations on a "frequency-trimmed" subset of verbs from historical corpora is the best approach to ensure data comparability (see Kodner 2019). Second,

contrary to Trips and Rainsford (2022), we advocate using corpus data in addition to lexicographical resources to address the attestation problem, showing that this prevents the analysis being affected by hapax constructions recorded in historical dictionaries which are very unlikely to have formed part of the learner's input. We conclude by suggesting a new template for researchers working with models of learnability in diachrony, in which a comparison with modern child-directed speech data forms an essential guide to the correct interpretation of the historical data.

References

- Dresher, B. Elan, and Aditi Lahiri. (2022) "The Foot in the History of English: Challenges to Metrical Coherence." In *English Historical Linguistics: Change in Structure and Meaning: Papers from the XXth ICEHL*, edited by Bettelou Los, Claire Cowie, Patrick Honeybone, and Graeme Trousdale, 42–59. Current Issues in Linguistic Theory, volume 358. Amsterdam ;Philadelphia: John Benjamins Publishing Company.
- Kodner, Jordan. (2019). "Estimating Child Linguistic Experience from Historical Corpora." *Glossa. A Journal of General Linguistics* 4, no. 1: 122.
- Kodner, Jordan. (2020). "Language Acquisition in the Past." Doctoral dissertation, University of Pennsylvania.
- Kodner, Jordan. (2022). "What Learning Latin Verbal Morphology Tells Us about Morphological Theory." *Natural Language and Linguistic Theory*. <https://doi.org/10.1007/s11049-022-09553-2>.
- Kroch, Anthony, and Ann Taylor, (eds). 2000. *The Penn-Helsinki Parsed Corpus of Middle English, Second Edition (PPCME2)*. Philadelphia: University of Pennsylvania.
- B. MacWhinney. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates; CHILDES -- NICHD HD082736.
- Ringe, Don, and Charles Yang. (2022). "The Threshold of Productivity and the 'Irregularization' of Verbs in Early Modern English." In *English Historical Linguistics: Change in Structure and Meaning: Papers from the XXth ICEHL*, edited by Bettelou Los, Claire Cowie, Patrick Honeybone, and Graeme Trousdale, 91–112. Current Issues in Linguistic Theory, volume 358. Amsterdam ; Philadelphia: John Benjamins Publishing Company.
- Trips, Carola, and Tom Rainsford. (2022). "Tolerating Subject-Experiencers? Yang's Tolerance Principle Applied to Psych Verbs under Contact in Middle English" Volume 6, Article 12: 1–43.
- Yang, Charles. (2016). *The Price of Linguistic Productivity*. Cambridge, MA: MIT Press.

**Marked vs. unmarked unaccusativity with alternating verbs:
Linking diachronic and experimental data.**

Alessia Cassarà¹, Achim Stein¹, Chantal van Dijk^{1,2}, Holger Hopp²

¹University of Stuttgart, ²TU Braunschweig

In this talk, we discuss how psycholinguistic studies can help in determining what constitutes a verb class and how this class changes over time. We focus on alternating verbs that can occur in transitive (causative) as well as intransitive (anticausative, unaccusative, inchoative) structures, such as Italian *rompere* ‘to break’. In languages like French or Italian, unaccusative verbs can be either be marked by a reflexive pronoun (R), as in *La chaise se casse* (‘the chair breaks’), or not (U), as in *La température augmente* (‘the temperature raises’). Although exhibiting different morphological properties, both marked and unmarked unaccusatives are said to share the same event structure, i.e. the absence of external causation and a subject with non-agentive properties (e.g., Dowty 1979, Jackendoff 1987, Levin & Rappaport Hovav 1995).

Many of the verbs which are marked with the reflexive pronoun in Modern French used to be unmarked in Old French (ex: *fondre* ‘to melt’ in Old French became *se fondre* ‘to melt’ in Modern French), suggesting that change has occurred in this verb class. Auxiliary selection is a second diagnostic that changed, e.g. from OF ‘be’ (*l’eve estoit refroidie* ‘the water has become cold’) to ModF ‘have’ (*l’eau avait refroidie*). We aim to determine what triggered this change from Old to Modern French. Specifically, we investigate whether their shared syntactic structure (presence or absence of the reflexive marker) or semantic factors (shared event structure) play a more decisive role.

We use psycholinguistic methods to address this question in experiments targeting these typical UA properties. In line with previous language processing and priming work by, e.g., Felser (2017) and Kootstra & Muysken (2019), we assume that instances of historical change can also be elicited in synchronic experimental conditions (“change in the lab”) and that the factors causing these changes should also elicit strong priming effects. In this sense, we believe that the use of psycholinguistic methods can shed light on the mechanisms underlying language change.

Several authors have pointed out that Old French displayed a situation similar to that found in Modern standard Italian (e.g. Gougenheim, 1973). In Old French as well as in Italian the diagnostics provide a well-defined definition of unaccusatives, as opposed to Modern French. Therefore we present two experiments on Italian, and plan similar experiments for Modern French in order to parallel the historical change from Old to Modern French.

In a first experiment, we tested the hypothesis that alternating verbs, when primed in their unaccusative form (e.g. *The vase broke*), elicit more unaccusative target structures than when primed with their transitive counterparts (e.g. *The wind broke the vase*). The results of a priming task with 48 native speakers of Italian show evidence of UA priming when primes and targets share the same verbs.

In an ongoing experiment, we explore whether UA priming obtains even if primes and targets contain different verbs. If we observe priming effects, they can be associated either with the syntactic (surface) structure (i.e. the presence or absence of a reflexive marker) or with the semantic (event) structure (i.e. the absence of external causation).

To disentangle the two factors, we designed a follow-up experiment, where the same production task is carried out with cross-lexical prime-target items involving different types of unaccusative verbs (reflexive marked unaccusatives, such as *la sedia si rompe* ‘the chair breaks’ vs. unmarked unaccusative verbs, such as *la pentola bolle* ‘the kettle boils’). Such an experiment will provide evidence as to whether syntactic overlap between prime and target (identical marking) leads to more priming than semantic overlap only (different marking). On the basis of these findings, we will discuss possible links between processing experiments and diachronic change. For instance, if the findings of the experiment reveal that syntactic structure has a stronger priming effect than event structure, we would expect change to occur with verbs that share the same surface structure (either reflexive-marked or unmarked). Conversely, if event structure exhibits a stronger effect, we expect change to have affected both unmarked and marked verbs simultaneously, by virtue of sharing the same semantic properties. More generally, we explore the hypothesis whether priming effects of event structure independent of syntactic overlap suggest that historical change occurs with reference to verb classes that are defined semantically (by shared event structure), rather than syntactically (by presence or absence of the reflexive marker).

References

- Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.
- Felser, Claudia. 2017. Syntactic Ambiguity in Real-Time Language Processing and Diachronic Change. In Hundt, Marianne & Mollin, Sandra & Pfenninger, Simone (eds.), *The Changing English Language: Psycholinguistic Perspectives*, 271–291. Cambridge: Cambridge University Press.
- Gougenheim, Georges. 1973. *Grammaire de la langue française du XVIe siècle*. Paris.
- Jackendoff, Ray. 1987. The Status of Thematic Relations in Linguistic Theory. *Linguistic Inquiry* 18. 369–411.
- Kootstra, Gerrit Jan & Muysken, Pieter. 2019. Structural Priming, Levels of Awareness, and Agency in Contact-Induced Language Change. *Languages*. 4(3). 65.
- Legendre, Géraldine. 1989. *Unaccusativity in French*. *Lingua* 79. 95–164.
- Levin, Beth & Rappaport Hovav, Malka. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.

Language Acquisition and a Process-Centered View of Language Change

Jordan Kodner
Stony Brook University

I argue that the actuation of a diverse range of diachronic phenomena in phonology, morphology, and syntax can be subsumed under the process of generalization learning during child language acquisition. These include a secondary split in 20th century Menominee and instance of phonemicization by phonological ‘rule reversal’ in Middle High German (Richter, 2021), the sporadic ‘irregularization’ of Early Modern English past tense forms (Ringe and Yang, 2022), the analogical extension of minority inflectional patterns at the expense of statistically predominant patterns in Late Latin past participles (Kodner, 2022) and Iranian Armenian aorists (Kodner and Dolatian, in prep), ‘Dative Sickness’ ongoing in Icelandic morphosyntax (Nowenstein et al., 2020), and the proliferation of the to-dative construction (Kodner, 2020) and argument structure change for psych-verbs (Trips and Rainsford, 2022) in Middle English. This has broad implications for how we conceptualize language change: an ontology of effects in language change will not line up with an ontology of processes. An approach to the study of change which focuses on *processes or mechanisms* over outcomes and effects stands to bring clarity to a confusing tangle of descriptive phenomena.

The model of generalization learning applied in these studies centers on the Tolerance Principle (TP; Yang, 2016), which provides an exact threshold for the number of exceptions that a linguistic generalization over some scope can tolerate if it is to be entered into a learner’s grammar. Over-regularizations, among the most common innovations in child productions (e.g., Xu and Pinker, 1995; Mayol, 2007) can result from a learner’s calculation over their limited linguistic experience: A TP calculation that would fail over an adult’s lexicon succeeds (perhaps transiently) for the learner, leading to innovation. It is applicable across generalization learning in phonology, morphology, and syntax because it separates the algorithmic aspect of acquisition from the representations over which generalizations are formed (Payne and Yang, 2023), thus a wide range of changes to the grammar may be subsumed under this single mechanism.

In every case investigated here, the TP calculated over acquisition-like samples (Nagy and Anderson, 1984; Yang, 2016; Kodner, 2019) from available corpora reveals patterns of (non-)productivity that are not evident from post-hoc statistical analysis. For example, the TP determines that the statistically predominant Latin participle patterns *-tus* and short *-itus* were actually unproductive. Indeed, they retracted or died out, consistent with this result. But, *-ūtus*, which often supplanted them in Romance, is calculated to be productive within its scope despite its rarity. Thus, this analogical extension works out quantitatively as a standard, albeit fortuitous, case of learner over-regularization. How an innovation like this progresses to language change requires additional population-level mechanisms:

Of course, individual childhood innovations do not entail population-level change, nor is every change child-driven (e.g., Labov, 1994, 2007; Stanford, 2015). Combining insights from competing grammars (Kroch, 1994), with the sociolinguistics of peer-oriented early childhood interaction (e.g., Roberts and Labov, 1995; Nardy et al., 2014; Loukatou et al., 2021), and experimentation on regularization and matching of variable input by children and adults (e.g., Hudson Kam and Newport, 2005; Newport, 2020; Austin et al., 2022), the quantitative predictions of the TP can be extended to model change in the face of population-level variation (Sneller et al., 2019; Kodner and Richter, 2020). This yields insights into why these innovations may progress through actuation and gain a foothold in a population while others may not. This in turn provides a means for distinguishing instances of child-driven from adult-driven change in cases where direct observation is no longer possible.

This work demonstrates that a single mechanism, over-generalization during language acquisition, unites several disparate effects ranging from cases of phonemicization to changes in argument structure. An approach to language change centering the mechanisms or processes (generalization learning, category learning, specific processes of phonetic perception (e.g., Ohala et al., 1981) and production, online syntactic processing, more broadly child- and adult-driven changes, etc.) reconceptualizes the problem space in a way that cross-cuts and reduces traditional taxonomies of effects (analogical leveling, extension, phonemicization, secondary splits, grammaticalization, bleaching, etc.) and opens the door for new insights into when, why, and how language change occurs.

References

- Austin, A. C., Schuler, K. D., Furlong, S., and Newport, E. (2022). Learning a language from inconsistent input: Regularization in child and adult learners. *LLD*, 18(3).
- Hudson Kam, C. L. and Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1(2):151–195.
- Kodner, J. (2019). Estimating child linguistic experience from historical corpora. *Glossa: a journal of general linguistics*, 4(1).
- Kodner, J. (2020). *Language Acquisition in the Past*. PhD thesis, University of Pennsylvania.
- Kodner, J. (2022). What learning Latin verbal morphology tells us about morphological theory. *Natural Language & Linguistic Theory*.
- Kodner, J. and Richter, C. (2020). Transparent /ai/-raising as a contact phenomenon. In Purse, R. and White, Y., editors, *PWPL: NWAV*, volume 25, pages 61–70.
- Kroch, A. (1994). Morphosyntactic variation. In *Proceedings of the thirtieth annual meeting of the Chicago Linguistic Society*, volume 2, pages 180–201.
- Labov, W. (1994). *Principles of linguistic change, volume 1: Internal factors*. John Wiley & Sons.
- Labov, W. (2007). Transmission and diffusion. *Language*, 83(2):344–387.
- Loukatou, G., Scaff, C., Demuth, K., Cristia, A., and Havron, N. (2021). Child-directed and overheard input from different speakers in two distinct cultures.
- Mayol, L. (2007). Acquisition of irregular patterns in Spanish verbal morphology. In *Proceedings of the twelfth ESSLLI Student Session*, pages 1–11, Dublin.
- Nagy, W. E. and Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, pages 304–330.
- Nardy, A., Chevrot, J.-P., and Barbu, S. (2014). Sociolinguistic convergence and social interactions within a group of preschoolers: A longitudinal study. *LVC*, 26(3).
- Newport, E. L. (2020). Children and adults as language learners: Rules, variation, and maturational change. *Topics in Cognitive Science*, 12(1):153–169.
- Nowenstein, I., Sigurjónsdóttir, S., Yang, C., Ingason, A., Wallenberg, J., et al. (2020). The meaning of case: Morphosyntactic bootstrapping and icelandic datives. In *45th Annual Boston University Conference on Language Development*. Cascadilla Press.
- Ohala, J. J., Masek, C. S., Hendrick, R. A., and Miller, M. F. (1981). The listener as a source of sound change. In *Proceedings of the Chicago Linguistics Society*.
- Payne, S. and Yang, C. (2023). Making good on BADS: Commentary on dressler et al. 2020. *Italian Journal of Linguistics*.
- Richter, C. (2021). *Alternation-Sensitive Phoneme Learning: Implications for Children's Development and Language Change*. PhD thesis, University of Pennsylvania.

- Ringe, D. and Yang, C. (2022). The threshold of productivity and the ‘irregularization’ of verbs in Early Modern English.
- Roberts, J. and Labov, W. (1995). Learning to talk Philadelphian: Acquisition of short a by preschool children. *Language Variation and Change*, 7(1):101–112.
- Sneller, B., Fruehwald, J., and Yang, C. (2019). Using the Tolerance Principle to predict phonological change. *Language Variation and Change*, 31(1):1–20.
- Stanford, J. N. (2015). Language acquisition and language change. *The Routledge Handbook of Historical Linguistics*, page 466.
- Trips, C. and Rainsford, T. (2022). Tolerating subject-experiencers? Yang’s Tolerance Principle applied to psych verbs under contact in Middle English. *JHS*, 6(12).
- Xu, F. and Pinker, S. (1995). Weird past tense forms. *Journal of Child Language*, 22(3):531–556.
- Yang, C. (2016). *The Price of Linguistic Productivity*. MIT Press, Cambridge, MA.