

Overgeneralization and change: The role of acquisition in diachrony

A commentary on

Cournane, A.: A developmental view on incrementation in language change

Marit Westergaard

UiT The Arctic University of Norway & NTNU Norwegian University of Science and Technology

This is a pre-print version reflecting changes made in the peer review and editing process, but it is not the publisher's PDF. This article appeared as:

Westergaard, Marit. 2019. Overgeneralization and change: The role of acquisition in diachrony. *Theoretical Linguistics* 45: 3-4, 225-231. Available [here](#).

When citing, please use the page numbers given there. The publisher should be contacted for permission to re-use or reprint the material in any form.

In a target article in *Theoretical Linguistics*, Cournane (2019) argues that child language acquisition is not only responsible for innovation in language change processes, but that young learners are also responsible for the incrementation of the change in a language community. She sketches an inverted U-shaped model of incrementation (p. 17), where children's initial overgeneralization of a linguistic phenomenon (input-divergence) is followed by a drop in production of the overgeneralized form (due to children's sensitivity to adult input), which is again followed by an increase in production in late childhood and adolescence (possibly caused by input from other children). The model is illustrated by cross-sectional data from different age groups in Toronto on two linguistic phenomena in English, modal interpretation and /u/-fronting (a change currently in progress in Toronto): For modality, Cournane (2015) and Cournane & Pérez-Leroux (under revision) have found that, although epistemic modality is acquired later than root modality, 5-year-olds display a clear preference for epistemic interpretations (of ambiguous sentences), thus overgeneralizing a newly learned property and crucially matching the shift from root to epistemic modality typically seen in diachronic development. For /u/-fronting, Hall & Maddeaux (2018) predicted that young children's production would correspond closely to the adult input while pre-adolescents would be more innovative, but instead they find that younger children are in fact the most different from the adults, while the older children's production is in between (and they thus sketch a development which they refer to as "two steps forward, one step back"). Based on these relatively sparse data, Cournane (2019: 18) concludes that "child generalization is in the direction of change ..., supporting the hypothesis that children's overgeneralizations are of the right kind to drive incrementation."

In general, I find this an interesting take on language change. It is especially welcome that the model addresses gaps in the literature by taking acquisition data from pre-adolescent children into account (not only from very young children) and that input from other children is also considered (not just from the parent generation). In my view, data from older children as well as the effect of their production on the development of younger children could very well be able to explain remaining puzzles in language acquisition and change. However, I would like to raise a question about the generalizability of the inverted U-shaped model and also provide some data from a change in progress that might be relevant to the discussion.

There seems to be one important restriction for the U-shaped model, in that Cournane indicates that it only applies to *some* linguistic phenomena (p. 1). This is further specified on p. 18, where she states that the model "is a simple descriptive sketch of the dynamics of learning over time relative to the target grammar, and it is not expected to apply to all learning, but at least to those kinds that rely on generalization and extension." Nevertheless, she also argues that "child learning is always about extending beyond finite input" (p. 2). While I agree that children always attempt to find the rules and systems behind the linguistic data that they are exposed to and in that process also make generalizations, most of them in small steps (micro-cues; see e.g. Westergaard 2014), I do not think it is correct to say that all language acquisition involves *over*generalization. Thus, if the model only applies to linguistic properties where we typically find that children overgeneralize the input, this will severely limit its scope and exclude a range of syntactic phenomena where young children are often found to be conservative learners (Snyder 2007) and to "undergeneralize" rather than overgeneralize (Roepers 1999: 175). That is, young children are often found to produce fewer overt elements than what is found in the input (see e.g. Valian 1990, 1991 for a comparison of null subjects in English and Italian child language) or less syntactic movement (see e.g. Westergaard 2014 on various constructions requiring movement in Norwegian; object shift, possessive movement, etc.), while the opposite is virtually never attested. For example, for

subject-auxiliary inversion in English questions, children have been found to occasionally produce sentences without movement, while overgeneralization of inversion is unattested, as it is not extended to contexts that don't require it (other clause types or verb types, see e.g. Radford 1992, Westergaard 2009a). Similar findings are attested in child language for verb second (V2) word order in other Germanic languages, e.g. Swiss German (Schönenberger 2001), Swedish (Waldmann 2011) or Norwegian (Westergaard 2009b). Since this type of undergeneralization – or preference for economy – has also often been attested in diachrony (see e.g. van Gelderen 2004), such historical developments have been analyzed as (partly) a result of child language acquisition, as I have also done myself, for the loss of V2 in the history of English (Westergaard 2009c, van Kemenade & Westergaard 2012) as well as for the development of non-V2 word order in *wh*-questions in many Norwegian dialects (Westergaard 2009d).¹ But the way I understand the inverted U-shaped model, historical development based on economy would not be accounted for by the model, since input-divergence is defined as overgeneralization.

Thus, for the inverted U-shaped model to have generalizable power, it seems to me that it is crucial that Cournane clearly defines which phenomena that the model is meant to account for. It is possible that this is to some extent an empirical question. Unfortunately, providing convincing evidence for the model seems to be a non-trivial task, as it would ideally require longitudinal data over many years, for phenomena that are either likely to undergo change or that are already in the process of changing. While longitudinal data would be most welcome in acquisition studies for a variety of reasons, collecting such data is time-consuming and expensive and therefore hardly feasible with normal resources for research.

But let us get back to overgeneralizations: Most extensions in child language seem to affect morphology, both in spontaneous production and in experimental work (e.g. Pinker 1999, Hudson Kam & Newport 2005). Morphology thus seems to be the most promising area to look for evidence for the inverted U-shaped model. I will therefore discuss some recent data on a change in progress in the Norwegian gender system that might be relevant in this respect.

Many Norwegian dialects are currently undergoing a rapid change from a three-gender system (masculine, feminine, neuter) to a two-gender system, in that feminine forms are lost and collapsed with masculine forms, which are by far the most frequent forms in the input.² From studies of child language corpora (e.g. Plunkett & Strömquist 1992, Rodina & Westergaard 2013), we know that young children overgeneralize masculine gender forms (especially the indefinite article) to both feminine and neuter nouns, as illustrated in (1)-(2), from Rodina & Westergaard (2013: 56).

- | | | |
|-----|--|---------------|
| (1) | <i>en mus</i>
a.M mouse.F.SG
Target: <i>ei mus</i> | (Emma 2;7.21) |
| (2) | <i>en hode</i> | (Ina 2;10.2) |

¹ Diachronic development is (of course) possible also in the other direction: The present-day strict V2 word order of German has been developed from a more variable system in Old and Middle High German (Hinterhölzl & van Kemenade 2012), and the non-V2 word order of a Northern dialect of Norwegian (originally the result of language contact with Kven/Finnish) is currently changing to V2 due to dialect leveling (Sollid 2005). In these cases, child language preferences are presumably overridden by other (sociolinguistic) factors; see also van Gelderen (2004).

² It should be noted that it is possible to use the two-gender system (common, neuter) in one of the written standards of Norwegian (*bokmål*), due to the historical connection with Danish. However, the three-gender system has been a robust property of the spoken language for centuries (until now, that is).

In order to investigate the ongoing change, Rodina & Westergaard (2015) used an elicited production experiment to test five different age groups in Tromsø (approximately 15 participants in each group). The distribution of feminine indefinite articles used with (previously) feminine nouns may be found on the left-hand side of Table 1.

Table 1: Results of two studies showing the development of feminine gender on indefinite articles in Norwegian (Tromsø) across five age groups of participants

	Rodina & Westergaard (2015)		Rodina & Westergaard (submitted)	
	Age	% of Fem	Age	% of Fem
Group 1	3;6-6;0	15%	3;7-5;9	1%
Group 2	6;6-8;2	9%	7;6-8;1	3%
Group 3	11;9-12;8	7%	11;10-12;9	29%
Group 4	18-19	56%	16	19%
Group 5	31-64	99%	32-67	92%

The finding that the youngest children numerically use somewhat more feminine forms than the older children in Groups 2 and 3 (15% vs. 9% and 7% respectively) corresponds to the original prediction made by Hall & Maddaux mentioned above, and Rodina & Westergaard also speculate that the production of the youngest children reflects the fact that their input mainly comes from adults, while older children are receiving more input from other children. However, it turns out that the difference between the three child groups is not statistically significant. Thus, the main picture that emerges from the statistical analysis is that the adults in Group 5 use the old feminine forms consistently, children (Groups 1-3) hardly use them at all, and 18-19-year-olds are in the middle. In principle, this gradual distribution of feminine forms could be used to argue for the inverted U-shaped model: The overgeneralization in young children is followed by retraction in older children, which still leaves them ahead of their input (the adult production). However, this is an unlikely scenario, mainly because the participants in Group 4 are simply too old, as retraction to the adult norm for the neuter nouns takes place already in the pre-school years (before age 6-7). Furthermore, the distribution within the group is that approximately half of the participants have a clear preference for the feminine, while the other half have a clear preference for the masculine. Another possibility for the inverted U-shaped model would be to analyze the child data collected here as the retraction phase and argue that at an (unattested) earlier stage, children are overgeneralizing even more. However, while overgeneralization of masculine forms is considerable in early child corpus data, there is no indication that it is ever 100% (Rodina & Westergaard 2013). Thus, it seems more likely that this distribution illustrates that this language change has started so recently and is happening so fast that it has not yet reached the adult generation (and has only partly affected the teenagers).

This experiment was replicated four years later in connection with another study (Rodina & Westergaard submitted), the only difference being that the teenagers in Group 4 are now approximately two years younger, thus corresponding more closely in age to participants belonging to Group 3 in the previous study. The results are found on the right-hand side of Table 1. The statistical analysis of these results shows that the only significant difference is found between the adults in Group 5 and all the other groups; that is, there is no longer a

statistical difference between Group 4 and the three child groups.³ This indicates that the change has now reached higher age groups, and importantly, that teenagers are not retracting from the overgeneralized form. Further evidence confirming this result is found in Busterud, Lohndal, Rodina & Westergaard (2019), which is a replication of the original Rodina & Westergaard (2015) study in Trondheim, where the change is even more advanced and presumably started earlier than in Tromsø. The distribution across the five age groups in the Trondheim study is 4%, 0%, 11%, 16% and 35%, and again, there is no statistically significant difference between Groups 1-4.

In conclusion, while the data from the development of Norwegian gender clearly show that this change is driven by children and adolescents, these studies do not provide any evidence for the inverted U-shaped model. However, as these studies are not longitudinal, the results do not constitute direct counterevidence for the model either. But they illustrate, in my view, how difficult it may be to find convincing evidence for it.

Acknowledgements: This research has been partly funded by a grant from the Research Council of Norway for the project MiMS (Microvariation in Multilingual Acquisition and Attrition Situations), project number 250857.

References

- Busterud, G., T. Lohndal, Y. Rodina & M. Westergaard. 2019. The loss of feminine gender in Norwegian: A dialect comparison. *Journal of Comparative Germanic Linguistics*.
- Cournane, A. 2015. *Modal development: input-divergent L1 acquisition in the direction of diachronic reanalysis*. PhD Thesis, University of Toronto (Canada).
- Cournane, A. & Pérez-Leroux, A. T. Under revision. Leaving obligations behind: epistemic incrementation in preschool English.
- Gelderen, Elly van. 2004. *Grammaticalization as economy*. John Benjamins.
- Hall, E. & Maddeaux, R. 2018. /u/-fronting and /æ/-raising in Toronto Families. Talk Presented at New Ways of Analyzing Variation (NWAV) 47. New York University.
- Hinterhölzl, R. & A. van Kemenade. 2012. The interaction between syntax, information structure and prosody in word order change. In T. Nevalainen & E. C. Traugott (eds.), *The Oxford Handbook of the History of English*. Oxford University Press.
- Hudson Kam, C. L. & E. 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.
- Kemenade, A. van & M. Westergaard. 2012. Syntax and information structure: Verb-second variation in Middle English. In A. M. Solin, M. J. López-Couso & B. Los (eds.) *Information Structure and Syntactic Change in the History of English [Oxford Studies in the History of English 2]*, 87-118. Oxford University Press.
- Pinker, S. 1999. *Words and Rules: The Ingredients of Language*. London: Weidenfeld and Nicolson.
- Plunkett, K. & S. Strömquist. 1992. The acquisition of Scandinavian languages. In D. I. Slobin (ed), *The crosslinguistic study of language acquisition*, 457–556. Lawrence Erlbaum Associates.
- Radford, Andrew. 1992. The acquisition of the morphosyntax of finite verbs in English. In J. M. Meisel (ed.), *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*, 23-62. Kluwer Academic Publishers.

³ The relatively high percentage of feminine forms in Group 3 is due to an almost exclusive preference for feminine in two of the 15 participants in this group and some variability in another two (the rest produce exclusively masculine forms in the relevant cases). For this reason, this group is not statistically different from Groups 1, 2 and 4.

- Rodina, Y. & M. Westergaard. 2013. The acquisition of gender and declension in a non-transparent system: monolinguals and bilinguals. *Studia Linguistica* 67.1, 47-67.
- Rodina, Y. & M. Westergaard. 2015. Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics* 27.2, 145-187.
- Rodina, Y. & M. Westergaard. Submitted. Grammatical gender and declension class in language change: A study of the loss of feminine gender in Norwegian. *Journal of Germanic Linguistics*.
- Roeper, T. 1999. Universal bilingualism. *Bilingualism: Language and Cognition* 2(3). 169–186. DOI: 10.1017/S1366728999000310
- Schönenberger, M. 2001. *Embedded V-to-C in Child Grammar: The Acquisition of Verb Placement in Swiss German*. [Studies in Theoretical Psycholinguistics]. Kluwer.
- Snyder, W. 2007. *Child language: The parametric approach*. Oxford: Oxford University Press.
- Sollid, H. 2005. *Språkdannelse og -stabilisering i møtet mellom kvensk og norsk*. Oslo: Novus.
- Valian, V. 1990. Null subjects: A problem for parameter-setting models of language acquisition. *Cognition* 35(2). 105–122. DOI: 10.1016/0010-0277(90)90011-8
- Valian, V. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40. 21–81. DOI: 10.1016/0010-0277(91)90046-7
- Waldmann, C. 2011. Moving in small steps towards verb second: A case study. *Nordic Journal of Linguistics*, 34(3), 331-359.
- Westergaard, M. 2009a. Usage-based vs. Rule-based Learning: The Acquisition of Word Order in Wh-Questions in English and Norwegian. *Journal of Child Language* 36.5, 1023-1051. DOI: 10.1017/S035000909009349.
- Westergaard, M. 2009b. The Acquisition of Word Order: Micro-cues, Information Structure and Economy. [*Linguistik Aktuell/Linguistics Today* 145], Amsterdam: John Benjamins.
- Westergaard, M. 2009c. The Development of Word Order in Old and Middle English: The Role of Information Structure and First Language Acquisition. *Diachronica* 26.1, 65-102. DOI: 10.1075/dia.26.1.03wes.
- Westergaard, M. 2009d. Microvariation as Diachrony: A View from Acquisition. *Journal of Comparative Germanic Linguistics*, 12.1, 49-79.
- Westergaard, M. 2014. Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation*, 14.1, 26-45.