# An elicited-production study of inflectional verb morphology in child Finnish

Sanna H. M. Räsänen,

University of Liverpool

Ben Ambridge & Julian M. Pine

University of Liverpool

ESRC International Centre for Language and Communicative Development (LuCiD)

RUNNING HEAD: FINNISH MORHOLOGY

ADDRESS FOR CORRESPONDENCE

Ben Ambridge, Psychological Sciences, University of Liverpool, Bedford Street South, Liverpool, United Kingdom. Email: Ben.Ambridge@Liverpool.ac.uk. Tel: +441517941109.

**Acknowledgments**

This research was funded by an ESRC PhD Scholarship awarded to the first author. We thank Viljami Venekoski for help with clarifying the measure of phonological similarity used.

**An elicited-production study of inflectional verb morphology in child Finnish**

Many generativist accounts (e.g., Wexler, 1998) argue for very early knowledge of inflection on the basis of very low rates of person/number marking errors in young children's speech. However, studies of Spanish (Aguado-Orea, 2004) and Brazilian Portuguese (Rubino & Pine, 1998) have revealed that these low overall error rates actually hide important differences across the verb paradigm. The present study investigated children's production of person/number marked verbs by eliciting present tense verb forms from 82 native Finnish-speaking children aged 2;2-4;8 years. Four main findings were observed: (1) Rates of person/number marking errors were higher in low frequency person/number contexts, even excluding children who showed no evidence of having learned the relevant morpheme, (2) most errors involved the use of higher-frequency forms in lower-frequency person/number contexts, (3) error rates were predicted not only by the frequency of person/number contexts (e.g., 3sg > 2pl), but also by the frequency of individual "ready-inflected" lexical target forms, and (4) for low-frequency verbs, lower error rates were observed for verbs with high phonological neighborhood density. It is concluded that any successful account of the development of verb inflection will need to incorporate both (a) rote-storage and retrieval of individual inflected forms and (b) phonological analogy across them.

**An elicited-production study of inflectional verb morphology in child Finnish**

An issue that lies at the heart of the cognitive sciences is the question of how children acquire their first language. The central theoretical debate in language acquisition research is between generativist theories, under which grammatical development involves the mapping of the target language onto innate grammatical rules, categories, principles and parameters (see Guasti, 2004; Lust, 2006; Crain & Thornton, 2012 for reviews), and constructivist theories (e.g. Bates & MacWhinney, 1989; Tomasello, 2000; 2003), which assume the gradual construction of a grammar on the basis of the language to which the child is exposed. Our goal in the present article is not only to pit these two approaches against one another in a domain that constitutes a particularly suitable test case – inflectional morphology – but to begin to move beyond this debate by identifying the processes that underlie *developmental changes* in children's use of inflections, and hence in language acquisition more generally.

One area that has proved useful as a testing ground for the debate between generativist and constructivist approaches to language acquisition more generally is children's acquisition of inflectional morphology (e.g., Berko; 1958; Cazden, 1968; Bowerman, 1973; Brown, 1973; Hoekstra & Hyams, 1998; Pine, Lieven & Rowland, 1998; Rasanen, Ambridge & Pine, 2014; Rispoli, Hadley & Holt, 2009; Theakston, Lieven & Tomasello, 2003; Wexler, 1998; Wilson, 2003). Since systems of inflectional morphology can be extremely complex (Finnish has approximately 260 verb inflections; Hakulinen et al., 2004), early error-free performance would appear to constitute evidence for innate abstract knowledge of inflection as posited by

generativist accounts. Constructivist accounts, in contrast, predict not only that children will make errors, but that the pattern of (in)correct use of inflections will directly reflect the input to which the child is exposed.

Thus the first goal of the present investigation of children's acquisition of Finnish verb morphology is to use this domain as a test case for the wider debate between generativist and constructivist approaches to morphology in particular, and to language acquisition in general. The second goal is to attempt to identify the causes of any observed developmental changes in children's proficiency with inflectional morphology; a goal that is all too often neglected in the cut and thrust of the debate between opposing theoretical positions. Again, our aim is not only to study morphological development for its own sake, but also to attempt to draw some conclusions about developmental changes in language acquisition more generally.

The structure of the remainder of this introduction is as follows. We begin by examining, in more detail, generativist and constructivist accounts of the acquisition of inflection and their predictions. Next we explore the extent to which these predictions have been supported by previous studies. Having briefly outlined the relevant properties of Finnish, we conclude by summarizing the design and predictions of the present study (including our analysis strategy for investigating developmental change).

First, a brief caveat is in order: Many readers will be familiar with the debate between single- and dual-route accounts of the English past-tense system (e.g., Pinker & Ullman, 2002; McClelland & Patterson, 2002). This debate concerns errors whereby children generate phonological forms that do not exist in the language (e.g., **sitted, *runned*), but use them in appropriate (past-tense) contexts. The issues

explored in the present study are orthogonal to this debate, since they concern errors whereby children produce phonological forms that *do* exist in the language (e.g., 3rd person singular verb forms), but use them in inappropriate contexts (e.g., 1st person singular contexts; analogous to errors such as **I sits* or **I runs* in English).

### *Generativist accounts of inflectional morphology and their predictions*

It is important at the outset to clarify our use of the term "generativist account" (Pinker, 1984; Harris & Wexler, 1996; Hoekstra & Hyams, 1998; Wexler, 1998; Deen, 2004; Legate & Yang, 2007). We include under this heading all accounts which assume that children begin the task of morphological acquisition with knowledge of (a) the functional category of INFLECTION (or AGREEMENT and TENSE), (b) the distinctions typically encoded by these categories (i.e., PERSON [1st/2nd/3rd; i.e., the speaker, listener and a third person respectively), NUMBER [singular/plural] and TENSE [past-present]) and (c) the syntactic category of VERB (as well as others that are less relevant for our present purposes; e.g., NOUN). These accounts assume, either implicitly or explicitly, that every verb form that bears PERSON/NUMBER AGREEMENT (and/or TENSE) marking is generated using a procedure that assigns or checks the relevant inflection. In other words, these accounts incorporate no significant role for rote storage of individual inflected forms. One possible exception is the generativist account of Pinker (1984), which would seem to allow for at least some rote storage; an issue to which we return in the discussion.

The technical details of these accounts are not important here (for a particularly clear exposition, see Blom and Wijnen, 2013: 227). The important point is the following: Because children are argued to begin the processes of morphological acquisition with a rule that assigns or checks the inflection of every agreement-marked (i.e., person/number marked) verb form, these accounts predict that – once the relevant inflections have been learned – children will never produce verb forms that bear incorrect person/number agreement marking (e.g., a 3sg form in a 1sg context[i]). Indeed, in each of the papers discussed above, this prediction is set out explicitly:

> Children simply don't say *I likes ice cream* [A 3sg form in a 1sg context]… The correct agreement features on verbal inflectional morphemes are known (Wexler, 1998: 42)
>
> Young German-speaking children… do not make agreement mistakes (Wexler, 1998: 19)
>
> A well established fact in child language is that errors of omission (e.g., *Mommy eat cake*) are extremely common, while errors of substitution (e.g., *I eats cake*) are very rare (Deen, 2004: 1).
>
> Errors of agreement [i.e., "substitution" rather than "omission" errors"; see above - BA] are superbly rare… These data strongly favor the analysis that children have an abstract rule of agreement at these early stages in development (Deen, 2004: 11)
>
> When finite forms are used, agreement is almost always correct (Hoekstra & Hyams, 1998: 84). [The caveat "when finite forms are used" reflects the widespread generativist assumption that, in many languages, TENSE/AGREEMENT marking is optional for young children (e.g., Wexler, 1998). However this consideration is not important for the present study, given that children rarely – if ever – omit TENSE/AGREEMENT marking in highly-inflected languages such as Finnish].
>
> Children's morphological errors…[do not reflect] use in inappropriate morphosyntactic contexts. (Legate & Yang, 2004: 322).
>
> The evidence we have adduced [for our hypothesis] includes…(c) Agreement is correct with main verbs (Harris & Wexler, 1996: 32)

As this last quotation makes particularly clear, a low rate of agreement-marking errors is presented not simply as a descriptive claim about children's language, but as a prediction of the relevant theories. Indeed, all take the finding that commission errors with person/number agreement marking are "rare (<1%)" (Rice, 2004:226), "vanishingly rare" (Wexler, 1998: 42), and occur at a rate that is "very low even by

the most stringent acquisition standards" (Hoekstra & Hyams, 1998: 84) as support for the claim that children have "Very Early Knowledge of Inflection" (Wexler, 1998) or show "Early Morphosyntactic Convergence" (Hoekstra & Hyams, 1998: 81).

As we have already seen, it is important to emphasize that generativist accounts only predict low error rates *provided that all of the relevant inflections have been learned* (e.g., Wexler, 1998: 42). Clearly, if a child uses (for example) a 3sg morpheme in a 3pl context, but only because she has yet to learn either (a) the phonological form of the 3pl morpheme or (b) that this phonological form *is* the 3pl morpheme, this cannot be taken as evidence against knowledge of an abstract system of inflection. Like a struggling second-language learner, the child could have abstract knowledge of the paradigm (i.e., know that she needs to add the 3pl inflection to the VERB stem), but not know what this inflection is. Consequently, when calculating error rates, it is important to include only data from children who have correctly produced a verb form that bears the relevant inflection in an appropriate context. (Given that generativist accounts take such correct productions as evidence for Very Early Knowledge of Inflection/Early Morphosyntactic Convergence [see quotations above], they cannot – at the same time – dismiss them as rote-learned forms that do not in fact demonstrate knowledge of the relevant morpheme and its person-number agreement properties).

Finally, it is important to note that all the generativist predictions outlined above – and tested in the present study - relate solely to errors of incorrect person/number agreement marking (e.g., the use of a 3sg verb form in a 1sg, 2sg or 3pl context). They do not relate, for example, to errors of tense omission (e.g., **Yesterday I play*) or – for languages that have different phonologically-based conjugation classes (e.g., Spanish, but not Finnish) - the use of one particular 3sg

inflectional morpheme in place of another. Hence, in order to be as generous as possible to generativist accounts, in the present study we treat as unscorable any verb form that is neither (a) correct nor (b) an unambiguous error of person/number agreement marking. That is, to be retained in the analysis, a response had to be either (a) correct or (b) an unambiguous error of person/number marking.

*Constructivist accounts of inflectional morphology and their predictions*

Constructivist accounts of morphological development (e.g., Bybee, 1995, 2001; Pizzuto & Caselli, 1992; Rubino & Pine, 1998; Pine, Lieven & Rowland, 1998; Gathercole, Sebastian & Soto, 1999; Aguado-Orea, 2004; Pine, Conti-Ramsden, Joseph, Lieven & Serratrice, 2008; Rasanen, Ambridge & Pine, 2014) assume that children do not start out with abstract categories of VERB, INFLECTION or AGREEMENT, and, instead, emphasize gradual, input-based learning. Children first store in memory complete, ready-inflected forms that they have heard used in the input (e.g., *halua-n* 'I want'). Initially, these chunks and frozen phrases function as unproductive rote-learnt forms, with the child being unaware of the internal morphological structure.

Only later in development do children generalize across these stored forms in a way that allows them to generate inflected forms of verbs that they have not heard in that particular form (including novel verbs in experimental studies). The precise characterization of this generalization process varies from theory to theory. Under *exemplar-based* models (see Skousen, Lonsdale & Parkinson, 2002, for a review), children store individual exemplars – i.e., ready-inflected verb forms – and generate

novel unattested forms ‘on the fly’, on the basis of phonological analogy to these stored forms. Other accounts (e.g., Janssen, Roelofs, & Levelt, 2002; Tomasello, 2003; Croft & Cruse, 2004; Boojj, 2010) posit morphological *schemas, constructions* or *slot-and-frame patterns* such as *[STEM]-n* (a putative 1sg schema in Finnish). However, it is unclear to what extent these accounts assume that morphological schemas are represented and stored independently in the brain, or use the term simply as a mnemonic for a particular type of exemplar-based generalization (e.g., Bybee, 2013). Accordingly, whilst the present article will make reference to “morphological schemas” we remain agnostic with regard to the issue of their independent representation.

It should be emphasized that whilst constructivist accounts assume that rote-learning plays a central role in the acquisition of verb morphology, they do not argue that *all* early knowledge of inflection consists of rote-learnt ready-inflected forms. Whilst this may be the case at the very earliest stages, the generalization processes outlined above are assumed to begin as soon as children have acquired a handful of stored forms. Thus, even children as young as 2 years (the youngest in the present study) are likely to have formed at least some productive schemas; in particular those for which the source forms are frequent in the input (e.g., 3sg *[STEM]-o*). On the other hand, even children as old as 5 years (the oldest in the present study) may have yet to form schemas for which the source forms are infrequent in their input (e.g., 2pl [STEM]-*tte*). Indeed, a study of novel noun marking in Polish (Dabrowska and Szczerbiński, 2006) found that even children aged 2;7 were highly productive (around 75% correct performance) with high-frequency inflections (e.g., masculine genitive), whilst children aged 4;5 showed poor performance for lower-frequency inflections (e.g., 15% for neuter dative).

How exactly does a child arrive at a correct person/number-marked verb form under constructivist accounts? First, the child searches memory for the appropriate stored ready-inflected form (token) for that verb. If none is found, the child will use one of the following strategies:

(a) Use a stored ready-inflected person/number-marked form that is available for direct recall from memory, either because it is of higher frequency than the target form – and so has a stronger representation in memory - or because another speaker has just produced it (e.g., Rubino & Pine, 1998). There is a trade-off here between availability and semantic/functional appropriateness (e.g., if the target is a 2pl form, it will usually be more appropriate to substitute a 2sg form [maintaining person] than a 3sg form [maintaining neither person nor number]).

(b) Generate the target form by phonological analogy with neighbours; stored forms that are phonologically similar and that bear appropriate person/tense number marking (e.g., Bybee, 1995; Marchman, 1997). For example, in Finnish, the 1sg present-tense form *kerää-n* '*I pick up'* might be generated by analogy with *herää-n* '*I wake up'*. Due to the highly regular nature of Finnish morphology, if an analogy with the target person/number-marked form is available, it will always yield the appropriate form (the same cannot be said for – for example – English irregular past-tense forms; Marchman, 1997). Under some versions of the account, this process could alternatively be conceptualized as retrieving a *[STEM]-n* morphological schema.

Thus, the predictions that follow from constructivist accounts are as follows:

(1) Although overall error rates may be relatively low, high error rates (and lower rates of correct use) will be observed for person/number contexts that are infrequent in the input and hence for which neither individual ready-inflected forms nor suitable morphological schemas are available in memory.

(2) Error rates will vary not only by person/number context, but also by target lexical form. Specifically, higher error rates (and lower rates of correct use) will be observed for target individual ready-inflected lexical verb forms (tokens) that are of low frequency in the input, and that are therefore represented only weakly – or not at all – in memory.

(3) Similarly, higher error rates (and lower rates of correct uses) will be observed for verbs with fewer phonological neighbors (i.e., with lower phonological neighborhood density), and hence fewer opportunities for successful phonological analogy. Since children are hypothesized to rely on phonological analogy only when a stored ready-inflected form is not available, constructivist accounts also predict an interaction such that phonological neighborhood density will have a greater effect for lower frequency than higher frequency lexical target forms. However, the importance of phonological neighbourhood density may decline into adulthood, as adults build the highly general representations that allow them to generate the semantically-appropriate person/number marked form for a verb, regardless of its phonological properties.

*Previous tests of generativist and constructivist predictions*

There is indeed some evidence to suggest that, as predicted by generativist accounts, children rarely produce person/number-marking errors. For example, Hoekstra and Hyams (1998) reviewed naturalistic data on overall rates of such errors in Spanish (Serra & Sole, 1992), Italian (Cipriani, Chilosi, Bottari & Pfanner, 1991; Pizzuto & Caselli, 1992), German (Clahsen & Penke, 1992) and Catalan (Serra & Sole, 1992). In all of the languages in the data reviewed, rates of person/number-marking error were very low (less than 5%). As noted above, these authors, as well as Wexler (1998) and Deen (2004), take these and similar findings as evidence for "very early knowledge of inflection" (Wexler, 1998: 25), and for innate knowledge of the abstract functional category of AGREEMENT (and TENSE).

However, there is some evidence from naturalistic studies of Spanish (Aguado-Orea, 2004) and Brazilian Portuguese (Rubino & Pine, 1998) that low overall error rates may hide important differences both across the verb paradigm - with higher error rates in lower frequency parts of the system – and across development. First, overall error rates are misleading because they collapse across data from both high and low frequency *person/number contexts* (or, from a constructivist viewpoint, *morphological schemas*). Rubino and Pine (1998) investigated naturalistic data from a child acquiring Brazilian Portuguese, and found that the overall rate of person/number marking errors was very low (3%). However, a closer look at the data revealed that this low error rate was composed of an error rate of 0.3% in high frequency 3sg contexts and of 43.5% error rate in low frequency 3pl contexts. Similar findings were reported by Aguado-Orea (2004) in a naturalistic corpus study of two Spanish-speaking children.

Second, overall error rates are misleading because they collapse across data from both high and low frequency *individual ready-inflected verb forms* that could in

principle be stored directly in the lexicon (e.g., Maratsos, 2000). For example, when Aguado-Orea (2004) removed just the two most frequent 1sg verb forms ("*I want*" and "*I can*") from the analyses, the error rate for 1sg contexts doubled from 4.9% to 10.4%.

Third, overall error rates are misleading because (presumably due to paucity of data) they tend to collapse data across long periods of time, ignoring the fact that the amount of data is likely to be unequal across different points in development. Given that children's rate of speech production generally increases with development, it is children's earliest speech, which is most likely to contain errors, that is generally under-represented.

Although these naturalistic studies would appear to provide some support for the constructivist prediction of high error-rates in low frequency parts of the system, they do not allow for investigation of the second and third constructivist predictions outlined above; that error rates will vary according to the frequency of the target lexical form and the phonological neighborhood density of the verb. This is simply because, in spontaneous speech, children (and, indeed, adults) tend to use only a small number of verbs, and – in most cases - only one or two inflectional forms of each (Aguado-Orea, 2004). The failure to test these predictions is an important omission, given that studies in other morphological domains have provided some evidence for the role of both lexical frequency and phonological neighborhood density (e.g., Marchman, 1997; Marchman et al., 1999; Dabrowska & Szczerbinski, 2006; Dabrowska, 2008; Kirjavainen, Nikolaev & Kidd, 2012).

Thus, the aim of the present study is to compare generativist and constructivist predictions regarding the development of inflectional morphology, using a method which allows for more control over the target verbs and inflectional contexts;

specifically elicited production, focusing on the Finnish present-tense system. Of course, we are by no means the first researchers to conduct an elicited-production study of verb morphology in a highly inflected language. Previous studies of this type include for instance Kunnari et al. (2011) in Finnish; Leonard, Caselli and Devescovi (2002) in Italian; Lukacs, Leonard, Kas and Pleh (2009) in Hungarian; and Stavrakaki and Clahsen (2009) in Greek. However, as far as we are aware, the present study is the most extensive of its type, with 1sg, 1pl, 2sg, 2pl and 3sg present tense forms elicited for each of 36 verbs, chosen to vary along the dimensions of lexical input frequency and phonological neighborhood density (defined in terms of morphophonological class size). Thus, to our knowledge, the present study constitutes the most comprehensive test to date of generativist and constructivist predictions regarding person/number-marking errors.

*Finnish*

An obvious advantage of testing these predictions in Finnish (a member of the Finno-Ugric group of languages, belonging to the Uralic family), is that Finnish is a highly inflected language. Finnish verbs (one popular dictionary, Hakulinen et al., 2004, lists approximately 9,000) mark both person and number, with six possible combinations: 1sg, 1pl, 2sg, 2pl, 3sg and 3pl (although, of course, verbs must agree with their subject, we use the term "person/number marking" as opposed to "subject-verb agreement marking", as overt subjects are rare in informal speech). Although Finnish is an agglutinative language, and sometimes includes a separate tense marker as well as a person/number inflection, this is not the case for the present tense, where only the

latter is used[ii]. An example of present tense verb conjugation is shown below using the verb *sano-a* '*to say'*.

| | |
|---|---|
| 1sg (minä) *sano-n* | 1pl (me) *sano-mme* |
| 2sg (sinä) *sano-t* | 2pl (te) *sano-tte* |
| 3sg (hän) *sano-o* | 3pl (he) *sano-vat* |

Unlike – for example – Spanish, Finnish does not have different conjugation classes. Thus, from the point of view of the adult linguist, a particular inflectional morpheme (e.g., 1sg *–n*) applies to all verbs. From the point of view of the child learning the system, however, the situation is far less straightforward. A complex system of morphophonological alternations involving vowel insertion, vowel harmony and consonant gradation[iii] means that the "same" inflection can be realized in many different ways, depending on the phonological properties of the verb. Indeed, the scheme adopted for the present study (see Appendix A) divides verbs into 20 morphophonological classes, each of which involves a different realization of any given tense/agreement marker (and more complex schemes propose as many as 46 classes).

Unlike English, Finnish verbs lack a free-standing, morphologically simple form: even the so-called *a*-infinitive, which corresponds to the English infinitive, has a separate inflection (e.g., *nous* + *ta* '*get up* + INF; *syö* + *dä* '*eat* + INF). However, it should be noted that, for some verbs, the infinitive is homophonous with the 3sg present tense form (see Appendix A)[iv]. In the present study, these ambiguous forms were scored as correct if they *could* have been correct (i.e., in 3sg contexts), but were otherwise excluded as unscorable, because we cannot tell whether children are

making a person/number marking error or instead producing an infinitive, which is a grammatical alternative for children under generativist "Optional Infinitive" accounts (e.g., Wexler, 1998).

A number of important considerations are in order with regard to colloquial spoken Finnish and its effects on verb morphology. First, in spoken speech, 3pl forms tend to be replaced by 3sg forms (e.g., Mielikäinen, 1984). Thus, it is perfectly acceptable to say, for instance, *Pojat juoksee* '*The boys runs'* instead of *Pojat juoksevat* '*The boys run*', even with an overt plural subject. For this reason, we did not elicit 3pl forms in the present study. Second, the passive form of the verb is generally used instead of the formal 1pl form in colloquial speech. For this reason, passive forms in 1pl contexts were counted as correct. Finally, 2pl forms can replace 2sg forms in formal contexts (like French *vous* forms). Because the study did not use formal contexts (children addressed a talking dog toy), such substitutions were treated as errors of person/number marking.

*Development*

As noted above, an important goal of the present study is not only to mediate between generativist and constructivist approaches, but also to begin to move beyond this debate by investigating the processes underlying any observed developmental changes in children's use of inflection (and – by extension – language in general). To this end, rather than following the more common approach of recruiting a number of different age groups, we instead tested a relatively large number of children (*N*=87) ranging over a wide age span (2;1-4;8). This approach allows us to study development by using statistical techniques that allow for the investigation of interactions between

continuous predictors (e.g., age in months and morphophonological class size). Thus if any observed development changes are underpinned by, for example, increasing use of phonological analogy with age, this phenomenon will surface as an interaction between these variables.

*Summary*

The present study compares the predictions of generativist and constructivist accounts of the acquisition of inflectional verb morphology by means of an elicited production study of Finnish present-tense inflection. Generativist accounts predict that, provided that the analysis is restricted to children who have learned the relevant person/number morpheme, error rates will be low across all inflectional contexts. Constructivist accounts predict low error rates for frequent contexts (e.g., 3sg), but higher error rates for low frequency (1) inflectional contexts and (2) individual lexical target forms. Constructivist accounts also predict (3) a negative correlation between phonological neighborhood density (i.e., morphophonological class size) and error rate and, perhaps, (4) a developmental decrease in the importance of phonological neighborhood density as learner's knowledge becomes more abstract, and hence less reliant on phonological analogy with close neighbours. Developmental changes in children's ability to supply correctly inflected forms are investigated by testing for interactions between these predictor variables and a continuous measure of children's age.

**Method**

*Participants*

There were 93 participants at the beginning of the study, recruited from six nurseries in Kuopio, Eastern Finland. All were typically developing, monolingual speakers of Finnish. No standardised language tests were used, but all the children were reported by their teachers and parents to exhibit typical language development. Eleven children were excluded because they did not attempt to respond on four consecutive trials. The final sample thus consisted of 82 participants (45 males, 37 females) with a mean age of 3;7 years (range 2;1-4;8).

*Design and materials*

The study employed a between-verbs, within-subjects design using an elicited production paradigm. The stimuli consisted of 36 verbs and accompanying videos, presented on a laptop computer. These verbs consisted of 18 high-frequency verbs and 18 semantically matched lower-frequency synonyms. The rationale behind selecting verbs in this way was to ensure a good spread of lexical target frequencies whilst minimizing, as far as possible, any confounding effect of semantics. Frequency counts (see below for details of how these were obtained) confirmed that each high frequency verb was indeed of higher frequency than its low frequency synonym and that, as a group, the former ($M$ = 26076, $SD$ = 29249) were significantly more frequent than the latter ($M$ = 2158, $SD$ = 4780), $t(17) = 3.59$, $p$ =.002). An important additional selection criterion for the target verbs was that they were easy to depict on video, and to act out with the child in the experimental setting.

The 36 verbs were divided randomly into two sets, each containing 9 high/low-frequency synonym pairs (with the constraint that very close phonological neighbors *lyödä* '*to hit*' and *syödä* '*to eat'* were not in the same set). Each child was randomly assigned to one of the two sets (the purpose of the sets was simply to reduce the number of trials that any one child had to complete). The same video was used for the high-frequency and low-frequency member of each synonym pair.

For each of the 18 verbs seen by a particular child, each of the following five target present-tense forms was elicited (for a total of 90 trials per child): 1sg, 1pl, 2sg, 2sg, 3sg (3pl forms were not elicited as these are usually replaced by 3sg forms in colloquial speech[v]).

*Predictor Variables*

Token frequency counts of each individual lexical verb form were obtained from the CSC Language Bank Newspaper corpora, which includes 131.4 million word tokens (www.csc.fi); the same corpus used in a previous study of Finnish past-tense inflection (Kirjavainen et al., 2009). Whilst it would, of course, have been preferable to use an electronic corpus of spoken language – ideally child-directed speech – no such corpus was available (though, as discussed in the Results section, a small paper-based corpus was used to verify counts of individual person/number marking contexts).

In order to check that the frequency counts obtained were representative of everyday spoken Finnish, we used an online rating task to obtain subjective frequency estimates from 50 native speakers (see Balota, Pilotti & Cortese, 2001, for evidence that such estimates are an excellent proxy for objective frequency counts). The

correlation between these frequency ratings and the counts from the newspaper corpus was high, suggesting that the latter provides a valid measure of lexical frequency.

As a measure of phonological neighbourhood density, the number of morphophonological classmates for each verb (see Appendix A for details) was taken from the *Ison Suomen Kieliopin Verkkoversio* (VISK; Hakulinen et al., 2004), generally considered to be the definitive reference grammar. However, even within this particular grammar, selecting a classification scheme is not straightforward, as there are various different ways to conceptualize similarity. The broadest scheme posits 6 major classes, grouping together all verbs that share a particular infinitival ending (e.g., *kisa**ta**, koho**ta** and hävi**tä***), ignoring differences between their inflected forms (e.g., *kis**aa**-n, koh**oa**-n and häv**iä**-n*). The disadvantage of using this scheme is that it assumes that learners are sensitive to phonological similarity at a highly abstract level (i.e., primarily at the level of the "transformation" between the stem and the inflected form [e.g., "*t*-drop"], rather than the inflected form itself: the form that children actually hear in the relevant contexts). Conversely, the most fine-grained scheme possible would posit different classes for forms that vary only in their application of language-general phonotactic changes such as consonant gradation and vowel harmony, and so would fail to capture similarities that are almost certainly psychologically real for speakers.

In an attempt to capture phonological similarity at level that is meaningful for our participants, we used the 20 minor classes of the VISK, collapsing five very similar classes into one. This results in a 16-class scheme (see Appendix A), though only 11 classes were represented in the present stimuli (many classes are very small and contain no verbs that would be suitable, given the context of the experimental "game"). Importantly, this scheme still conceptualizes similarity in terms of the

inflected forms that children hear in the relevant contexts (e.g., *kis**aa**-n, koh**oa**-n* and *häv**iä**-n* each belong to a separate class, rather than a single "*t*-drop" class).

*Procedure*

Each child was tested individually in a quiet setting, with each session lasting approximately 15-25 minutes, depending on the child. Trials were presented in random order. Videos were shown on a laptop computer (13 inch screen). Audio recordings of the experimental sessions were made using Audacity 1.3.13 (running in the background on the same laptop).

The child was seated in front of the laptop computer, with the "talking" toy dog positioned so that it was behind the laptop and could not therefore "see" the laptop screen, but faced towards the child and the experimenter. The toy dog's internal speakers were connected to the laptop. First, the child completed a brief warm-up that involved being introduced to the toy dog and the experimenter. The child was told that he or she would be playing a game with the experimenter in which they would watch some videos of the experimenter and the toy dog acting out some actions together, and they would also be performing the actions. The child was told that her task would be to help the toy dog out by answering its questions. The experimenter then brought up the first video, and told the child, for example, that *Tässä on leikkaamista* [*This is cutting*]. Thus, the children were given the target verb in the form of a verbal noun in the partitive. This form was used because it has already undergone the "changes" that must be made to an infinitive form before the "addition" of the appropriate person/number morpheme (i.e., it contains the inflectional stem rather than the infinitival stem). Consequently, the task facing the

children is simpler than it would have been had the verb been presented in infinitival form. Throughout the experiment, if the child had trouble recalling the target verb, the experimenter repeated the target verb in this form. If the child used a non-target verb, that trial was classified as unscorable.

The questions asked by the dog varied according to the target form being elicited. For instance, for 2sg forms, the toy dog asked *Mitä minä teen*? [*What am I doing?*], while the child watched a video of the dog performing the relevant action. For 1pl forms, the experimenter and child performed the relevant action, while the dog asked *Mitä te teette*? [*What are you-pl doing?*]. The question probes for each target inflection are given in Table 1. Each video lasted for 5-6 seconds, and was played continuously during each verb trial to emphasize the ongoing nature of the action, and thus to encourage the use of the simple present tense form (Finnish has no present progressive), rather than, for example, the past tense. As an incentive, children were rewarded with stickers throughout the experiment, regardless of the responses produced.

INSERT TABLE 1 ABOUT HERE

*Transcription, coding, and reliability*

Responses were transcribed from the audio recordings and coded by the first author. The total number of responses was 7380 (5 target forms x 18 verbs x 82 participants). Responses were coded as (1) correct, (2) incorrect or (3) unscorable, as described below.

*(1) Correct inflection* (*N*=4343): The child used the correct person/number marked form of the appropriate verb, given the target context (because subject omission is very common, it was necessary to score relative to the target context, as opposed to the subject).

*(2) Incorrect inflection* (*N*=717): The child produced a person/number marked form of the appropriate verb, but one that was not appropriate given the target context.

*(3) Unscorable* (*N*=2320). The child produced a) no response or an unintelligible response (N=1350), b) a repetition of the dog's question (N=198), c) a non-present-tense form of the target verb (e.g., stem or infinitive) (N=101), or c) any form of a non-target verb (N=671). Although the proportion of unscorable responses (31.44%) is relatively high, many of these errors constitute pragmatically appropriate responses to the description task, and are thus very difficult to pre-empt entirely.

In some respects, whether a particular response counts as "incorrect" versus "unscorable" depends on the theoretical stance taken. Given our own theoretical position, our goal in classifying responses as incorrect versus unscorable was to be as generous as possible to generativist accounts, and as strict as possible with regard to constructivist accounts. Thus, we followed Harris and Wexler (1996), Hoekstra and Hyams (1998), Wexler (1998), Deen (2004) and Legate and Yang (2007) in counting as "incorrect" only incorrectly person/number-marked forms of the target verb. Given that other non-target responses are difficult to interpret, including such responses as incorrect (rather than unscorable) would have artificially inflated the error rate, which is predicted by generativist accounts to be very low.

By the same token, since the constructivist account predicts that children may use evasion strategies for low frequency, unfamiliar items, our decision to count any

possible instances of evasion as unscorable rather than incorrect biases the analysis against the constructivist position. Indeed, an ANOVA ($F(4,7220)=7.07$, $p < .001$) revealed that unscorable responses were less frequent for 3sg targets (always the most frequent input form) than 1sg, 1pl, 2sg and 2pl targets ($p =.006$; $p < .001$; $p = 0.007$; $p < .001$, respectively). Thus by excluding such responses from the analysis, we are minimizing the likelihood of observing frequency effects, and hence providing for a relatively conservative test of the constructivist claim that error rates are related to the frequency distribution of forms in the input.

The effect of these missing data should not be overstated, however. On average, a scorable response for each verb was contributed by 31 of the 41 children tested (SD=7.8). Furthermore, the rate of unscorable responses decreased significantly with age (simple Pearson $r=.396$ $p < .001$); a finding which provides some reassurance that missing data was largely a consequence of memory and processing limitations.

As the focus of the present study was children's correct and incorrect use of person/number marking, phonological errors involving the *verb stem only* were ignored. Again, the rationale behind this decision was to be as generous as possible to generativist accounts, by counting as correct any response in which the child is clearly attempting to produce the target person/number marked inflection. This decision biases the analysis against constructivist accounts, which would predict higher rates of such errors for target forms that are of low frequency and/or phonological neighbourhood density. An analysis revealed that children did indeed make more stem errors when the token frequency was lower and when the syllable length was longer ($\beta = -0.01$, $SE = 0.001$, $z = -2.14$, $p = .032$ and $\beta = 1.40$, $SE = 0.57$, $z = 2.46$, $p = .014$, respectively)[vi].

Thus the verb was considered to be the target verb if the stem included (a) a gradation error (e.g., *nou**se**taan* instead of *nou**st**aan*, (b) a local dialect form (e.g., *luk**k**ee* instead of *lukee*; *sy**yv**ään* instead of *sy**öd**ään*), c) misarticulations of consonants (e.g., *kä**l**elette* instead of *kävelette*) or (d) other modifications that still represented clear attempts at the target form (e.g., shortenings, such as *myhäämme* instead of *myhäilemme*). In order to calculate reliabilities, 10% of the responses were transcribed independently by another native Finnish speaker blind to the hypotheses under investigation. Agreement was 97.6%. Any disagreements were subjected to re-listening until agreement was reached.

**Results**

Because the constructivist approach predicts differences in error rates across different target inflectional contexts and across different verbs, in what follows, we generally report error rates by items rather than by subjects (the generativist prediction of very low error rates applies either way). On the more-usual by-subjects calculation, rates of correct use and error were 85.83% (SD=34.88%) and 14.17% (SD=34.90%) respectively (excluding unscorable/ambiguous/infinitival forms from the denominator). The mean proportion of correct inflections for each verb, collapsing across all inflectional contexts, is displayed in Table 2 (again, unscorable/ambiguous/infinitival forms were excluded from the denominator). Overall, children's performance was relatively good, with 85.83% correct performance.

Thus, whether the data are analysed by subjects or by items, it is clear that, on trials where they attempted to produce a present-tense form of the relevant verb,

children appeared to understand which person/number form was the target in each experimental scenario. This is important, as children very rarely produced subjects (as is usual in Finnish for 1st and 2nd person forms in general, and for 3rd person forms when the referent has already been established [here, by the dog's question]). Stem-only errors (N=32, plus N=35 errors that are ambiguous between stems and 3sg forms) and infinitive errors (N=20, plus N=14 errors that are ambiguous between infinitives and 3sg forms) were rare (and were counted as unscorable).

INSERT TABLE 2 ABOUT HERE

*Analysis by target inflectional context*

The overall rate of person/number-marking errors observed was 14.17%. Whilst this error rate is already somewhat higher than rates typically taken as evidence for virtually error-free performance (around 5%; Hoekstra & Hyams, 1998; Wexler, 1998), it hides considerably higher rates in certain parts of the system. Table 3 shows error rates broken down by target inflectional context (again excluding unscorable and ambiguous responses as outlined above). The pattern is very similar to that observed by Aguado-Orea (2004) and Rubino and Pine (1998), with a very low error rate for 3sg forms (<1%) hiding rates as high as 32% in other contexts.

INSERT TABLE 3 ABOUT HERE

However, as we saw in the introduction, generativist accounts predict low error rates only from the point at which children have learned the relevant inflection. We therefore recalculated these error rates, excluding – for each person/number context separately - children who did not produce at least one correct target form (see Table 3). For no inflectional context did this involve excluding more than 20% of children. Perhaps surprisingly, this made very little difference to the error rates, with rates as high as 32% observed. The finding that non-3sg contexts displayed error rates of 10%, 12%, 14% and 32% - even when controlling for knowledge of the relevant inflection – does not sit comfortably with the generativist prediction of "vanishingly rare" errors (Wexler, 1998: 42). Although it is not clear exactly what constitutes a "very low" error rate (Hoekstra & Hyams, 1998: 84), if rates of <5% are to be taken as evidence for this claim, it would seem inconsistent to argue that a rate that is higher by a factor of 7 does not constitute evidence against it.

Of course, as noted by an anonymous reviewer, some of these errors may have a pragmatic element. For example, if the child is asked by the dog "*What are we [the dog and the experimenter] doing?*", and responds with a 2sg form rather than a 2pl form, this could be a pragmatic rather than morphological error, or indeed not an "error" at all; the child may simply prefer to describe the actions of the dog alone, rather than of the dog and the experimenter together (remember that children almost never provided overt subjects). Note, however, that by excluding data from children who did not produce at least one instance of the target inflection, we are restricting the analysis to children who not only clearly understood which form they were supposed to be producing in each person/number context – but were also willing and able to do so. That said, it is probably impossible to design an experimental task that rules out this objection altogether; ultimately only a speaker can decide who she will

address, or whose actions she will describe (an issue to which we return in the discussion)

A further justification for including such responses as errors is that doing so biases the analysis against observing the effects predicted by the constructivist account. If these really are pragmatic errors – or not errors at all – there is no reason to expect them to pattern by target context, lexical frequency or phonological neighbourhood density of the target form. Of course, including such forms as errors also drives up the overall error rate, at the expense of the generativist account. But any finding that errors pattern according to these factors would support the constructivist over the generativist account in any case, regardless of the overall error rate.

An important point to note with regard to these person/number marking errors, and their implications for generativist accounts is that the observed error rates are not only high, but also uneven (see Table 3). A one-way ANOVA revealed a significant main effect of target inflectional context, $F(4,5055)=157.46$, $p< .001$. Post hoc tests revealed that 2pl contexts - the least frequent in the corpus (see Appendix B) - attracted significantly more errors than all other contexts ($p< .001$ for all comparisons). Conversely, 3sg contexts – the most frequent in the corpus (see Appendix B) – attracted significantly fewer errors than all other contexts ($p< .001$ for all comparisons). Children also produced significantly more errors in 2sg than 1sg contexts ($p = .045$), with no other significant differences observed.

Recall that the frequency ranking of contexts discussed above is based on a newspaper corpus. In order to check that error rates were higher for person+number contexts that are of low frequency in speech to children, we calculated proportions of

different present tense forms in 17 short paper-based transcriptions of child-adult interactions (total length 678 minutes), made available by the University of Oulu in Finland (it was not possible to use this corpus for the main analysis, as the majority of the verbs used in the present study did not appear at all in this relatively small corpus). This analysis was done by hand. All verbs in the corpus were included, regardless of whether or not they appeared in the present study, except for the extremely frequent verb *olla, 'to be'*. Fig. 1 shows the relationship between these counts and the proportion of person/number errors for each inflectional context in the present study. The claims above regarding frequency of individual contexts (3sg most frequent, 2pl least frequent) were clearly supported (though the correlation - simple Pearson $r$=-.785 - was not significant due to the small sample size: $N$=5).

INSERT FIGURE 1 ABOUT HERE

The final important point to note from this analysis is that many errors involve the substitution of a higher-frequency form for a low-frequency target form (see Table 4). For example, 42% of errors were substitutions of more frequent forms (mostly 2sg, 1pl passive, or 3sg) for 2pl forms; the least frequent in both the adult and child corpora (see Rasanen et al., 2014, for evidence of "defaulting" to high frequency forms in English). In contrast, fewer than 1% of errors were substitutions of less frequent forms for 3sg forms (the most frequent in both corpora).

INSERT TABLE 4 ABOUT HERE

The findings that (a) errors are more frequent for low frequency target contexts and (b) almost always involve replacement by higher-frequency forms are clearly consistent with constructivist approaches, which emphasize the importance of frequency-sensitive input-based learning. However, a stronger prediction of such approaches is that errors will pattern according to properties of the input distribution at the level of individual verbs. The following analyses test this prediction, using a developmental approach designed to elucidate the processes underlying changes in children's use of inflectional morphology.

*By-verbs Analysis*

The analysis reported above compared the generativist prediction of low overall error rates against the constructivist prediction of high error rates for low frequency *target contexts* (e.g., 2pl vs 3sg). In order to test the second and third constructivist predictions outlined in the introduction – that error rates will be lower for (a) high frequency lexical target forms and (b) verbs with high phonological neighborhood density – a finer-grained *by-verbs* analysis is required.

In order to examine patterns of correct use versus error across all of the 180 different target forms elicited in the study (36 verbs x 5 person/number contexts) we constructed mixed-effects regression models with items and participants as random effects (see Baayen, 2008). Compared with traditional by-subjects/items regression analysis, the advantage of using such an approach is that mixed-effects modeling takes into account both by-subject and by-item variation, and thus is more powerful. As the outcome measure was dichotomous (for each target, each child produced either a correct or an incorrect form [coded as 1/0], with all other responses, including bare

stems, infinitives and ambiguous forms, treated as missing data), binomial logistic regression models were used. The fixed effects of interest were the input token frequency of the target lexical verb form (e.g., *sano-n, sano-t, sano-o, sano-mme, sano-tte;* taken from the newspaper corpus, as most did not occur in the small child-directed corpus discussed above), morphophonological class size and age. Verb length (in syllables) was included as a control predictor: Under any theoretical account, longer verbs might be expected to introduce more processing difficulty and hence increase error rates. Verb set (A or B) was not included, as preliminary analyses revealed that it was not a significant predictor of rates of correct production ($\beta$ = -0.21, *SE* = 0.22, $z$ = -0.96, $p$ = .337). All model comparisons used likelihood ratio tests performed in R with the anova function. The details of all statistical models are presented in Table 5.

INSERT TABLE 5 ABOUT HERE

Model 1 (a reduced, baseline model) included only verb length (in syllables) as the (control) predictor variable. Verb length was a significant negative predictor of children's ability to supply the correctly inflected target form ($\beta$ = -0.35, *SE* = 0.11, $z$ = -3.22, $p$ = .001).

Model 2 added the predictor of lexical verb form token frequency. Whilst verb length was no longer a significant predictor ($\beta$ = -0.15, *SE* = 0.12, $z$ = -1.25, $p$ = .211), token frequency was a large and significant positive predictor of children's ability to supply the correctly inflected target form ($\beta$ = 0.04, *SE* = 0.007, $z$ = 6.86, $p$ < .001); i.e., a negative predictor of the error rate. This model (AIC = 3778, logLik =-1884) provided a significantly better fit to the data then the reduced model (AIC = 3880,

logLik =-1936; $p < .001$). This finding provides support for the constructivist claim that an important mechanism in early morphological development is the storage and retrieval of ready-inflected forms, and is problematic for those generativist accounts (e.g., Hoekstra & Hyams, 1998; Wexler, 1998) that seek to explain person/number marking errors solely in terms of children's lack of knowledge of particular inflectional *morphemes*.

Model 3 added the predictor of morphophonological class size. This variable did not predict the rate of correctly inflected target forms ($\beta$ = -0.01, $SE$ = 0.067 $z$ = -0.12, $p$ = .907), and the model (AIC = 3780, logLik =-1884) did not provide a significantly better fit to the data than Model 2 ($p$ = .901). However, this finding needs to be interpreted in the context of a significant interaction between token frequency and morphophonological class size observed in Model 4 ($\beta$ = -0.03, $SE$ = 0.007, $z$ = -4.38, $p < .001$), which provided a significantly better fit to the data than Model 2 (AIC = 3760, logLik =-1873) ($p < 0.001$). The interaction is plotted in Fig. 2, from which it can be seen that morphophonological class size had a larger facilitative effect for lower frequency than higher frequency target verb forms. The direction of this interaction suggests that, consistent with the constructivist approach, children rely on phonological analogy only when a stored ready-inflected form is not available.

INSERT FIGURE 2 HERE

*Development*

In summary, the findings outlined above – a non-negligible rate of errors, that pattern according to (a) person/number context, (b) target lexical form frequency and (c) morphophonological class size (for lower frequency verbs) – would appear to sit more

comfortably with constructivist than generativist approaches. As noted in the introduction, however, an important goal of the present work was to begin to move beyond this debate, and investigate in more detail the factors that appear to influence development in children's use of inflection.

To this end, Model 5 added the children's age (in months) and its two- and three-way interactions with lexical verb form token frequency and morphophonological class size. This model provided a significantly better fit to the data than Model 4 (AIC = 3748, logLik =-1863) ($p < .001$). A main effect of age was observed ($\beta = 6.69$, $SE = 1.57$, $z = 4.27$, $p < .001$), reflecting the fact that, as would be expected under any account, older children are better at supplying the correct target form.

More interestingly, this analysis revealed a significant interaction between age and morphophonological class size ($\beta = -1.63$, $SE = 0.72$, $z = -2.28$, $p = .023$), but no other main effects or interactions. The negative interaction between age and morphophonological class size, shown in Fig. 3, reflects a decrease in the importance of morphophonological class size with age; a finding that is presumably due to learners' knowledge of the system becoming increasingly abstract with age, leaving them less reliant on analogy with close phonological neighbours.

INSERT FIGURE 3 HERE

Under a strict exemplar-based approach, this "more abstract knowledge" would constitute (a) more stored exemplars of each person+number-marked verb form and (b) the ability to analogize across all stored forms with the appropriate person+number features, presumably on the basis of this shared function, even in the

absence of close phonological similarity. Under an account that posits the independent representation of linguistic generalizations, this "more abstract knowledge" would constitute stored *morphological schemas, constructions* or *slot-and-frame patterns* such as *[STEM]-n* (a putative 1sg schema), with the [STEM] slot having no particular phonological restrictions (due to the phonological heterogeneity of the verb forms in the input that gave rise to this schema).

Either way, the developmental pattern is clear: At all ages, children make use of rote-learned individual ready-inflected verb forms (hence the main effect of lexical target form frequency and no interaction with age). At all ages, children make use of phonological analogy with stored forms, when the target form is of low frequency in the input, and therefore not stored (hence the negative interaction of frequency and morphophonological class size, but no three-way interaction with age). However, as development proceeds, children become less reliant on phonological analogy with stored forms, as they gain the abstract knowledge needed to supply the correct person/number form of any verb, regardless of its phonology (hence the negative interaction of age and morphological class size).

**Discussion**

The present study constituted an elicited production study of Finnish present tense verb forms, designed to test the predictions of generativist and constructivist accounts of the acquisition of inflectional morphology. Four main findings were observed.

The first is that rates of person/number marking errors were as high as 32% for low frequency person/number contexts, even when excluding data from children who showed no evidence of having learned the relevant morpheme. This finding is

predicted by constructivist accounts, but is more difficult to reconcile with generativist accounts, which predict very low error rates (at least, once the child has learned the relevant morpheme). Note that because we excluded children who did not produce at least one correct instance of the relevant person/number morpheme, this pattern cannot solely be a reflection of a tendency to avoid using certain person/number contexts for pragmatic reasons.

The second is that most errors involved the use of higher-frequency forms in lower frequency person/number contexts. Again, this finding is more consistent with constructivist accounts, but could, in principle, be reconciled with generativist accounts, provided that one is prepared to additionally posit a significant degree of rote-use of high-frequency lexical target forms, even after the point at which individual person/number marking morphemes appear to have been acquired; though it is important to stress that none of the generativist accounts discussed thus far do so.

The third is that error rates were predicted not only by the frequency of person/number contexts (e.g., 3sg > 2pl), but also by the frequency of individual "ready-inflected" lexical target forms. Again, this finding is predicted by constructivist accounts, which posit an important role for rote-learning of individual lexical forms, and could be explained by a generativist account that adopted this assumption. As we noted in the introduction, the generativist accounts discussed up to this point implicitly rule out this assumption by taking all correctly inflected forms – even high frequency forms that could be rote learned - as evidence of abstract knowledge of inflection.

The fourth finding is that – for low frequency verbs - lower error rates were observed for verbs with high phonological neighborhood density, which allows children to generate otherwise-unavailable target forms by phonological analogy with

stored neighbours. However, a negative interaction of age and morphological class size indicated that, as development proceeds, children become less reliant on phonological analogy with stored forms, as they gain the abstract knowledge needed to supply the correct person/number form of any verb, regardless of its phonology. Again, these findings are consistent with constructivist accounts, which posit a role for phonological analogy with stored neighbours, "regular" or "irregular" alike (e.g., Ambridge, 2010). None of the generativist accounts discussed up to this point incorporate a rule for phonological analogy in regular systems.

Do our findings therefore count against only the *particular* generativist accounts discussed in the introduction (Harris & Wexler, 1996; Hoekstra & Hyams, 1998; Wexler, 1998; Deen, 2004; Legate and Yang, 2007) or against generativist approaches to morphological acquisition *in general*? On the one hand, it is certainly true that these data count most straightforwardly against those particular accounts, which specifically and explicitly predict low rates of person/number marking error.

On the other hand, one could, in principle, posit a generativist account that assumed - in addition to early knowledge of an abstract person/number marking system - both (a) considerable use of rote-learned ready-inflected verb forms for a protracted period and (b) phonological analogy across such forms, even for regular forms. However, since such an account would, in effect, constitute a constructivist account with innate knowledge of an abstract system of verb inflection added on, the onus would be on the proponents of such an account to explain exactly what explanatory power the additional innate knowledge is adding. In particular, note that such an account could not take low error rates as evidence for innate abstract knowledge, as low error rates (i.e., high rates of correct use) could reflect the use of rote-learned forms.

One generativist account that exhibits some of these characteristics is Pinker's (1984) paradigm-building account. While this account shares with other generativist accounts the assumption that children start out with abstract knowledge of the cells of person/number marking paradigms, the process by which children fill in these cells – effectively generalizing gradually across stored exemplars – has more in common with constructivist accounts. We suggest, however, that the present findings nevertheless constitute evidence against Pinker's (1984) proposal, for two reasons. The first is the reason that we gave above: Given the present evidence that children are storing and gradually generalizing across individual inflected forms (as assumed by Pinker's, 1984, account), additionally positing innate abstract knowledge of the paradigm would seem to add little or nothing to the explanation.

The second reason is that, in order to account for the phonological neighbourhood effects observed in the present study, Pinker's (1984) account would have to add the assumption of phonological analogy across all stored forms. However, in his work on another domain of inflectional morphology - the English past-tense - Pinker explicitly rules out such a mechanism. For example, although Prasada and Pinker (1993) argue for phonological analogy across stored *irregular* verb forms, they not only argue specifically against the possibility of phonological analogy across *regular* morphological forms, but present a study designed specifically to provide empirical evidence against this possibility (whether or not it successfully does so is a matter of debate; see Albright & Hayes, 2003; Ambridge, 2010). Since all of the verb forms in the present study are regular, Prasada and Pinker's (1993) account would seem to specifically predict that an effect of phonological neighborhood density will *not* occur.

Thus, although no study could ever provide definitive evidence against all possible future generativist accounts, on our view, the present findings both (a) constitute evidence against all generativist accounts that have been proposed so far (including Pinker, 1984) and (b) suggest that any future account would have to include such a large role for exemplar storage and analogy, that its generativist underpinnings would be seriously undermined: Given that children generate non-rote-stored forms by analogizing across stored exemplars, what do we gain by positing that – sometimes – they may additionally generate them using algebraic rules based on an innate abstract system?

It must be noted that due to the nature of Finnish, an obvious limitation of the present study is the fact that it is impossible to know for certain that, when apparent errors were made, children were indeed attempting to produce the target person+number marked form (subjects are almost always omitted in conversational Finnish). Note, however, that because we excluded children who did not produce at least one correct instance of the relevant person/number morpheme, the high error rates observed for certain inflectional contexts cannot solely be a reflection of a tendency to avoid these contexts for pragmatic reasons (e.g., using a 2sg form rather than a 2pl form because the child prefers to describe the actions of the dog alone, rather than the dog and the experimenter). The fact that these errors (e.g., substitutions of 2sg for 2pl) did not occur at random, but were predicted by both token lexical frequency and phonological neighborhood density provides further evidence that at least the majority were indeed errors, rather than pragmatic substitutions. Furthermore, the pattern of results is very similar to that observed in naturalistic studies of Spanish (Aguado-Orea, 2004) and Brazilian-Portuguese (Rubino & Pine, 1998), suggesting that any occasional misclassifications of errors as well-formed

attempts at non-target person+number forms did not substantially affect the overall pattern observed.

Nevertheless, in order to clarify this issue, it may be useful to conduct future studies using paradigms that encourage the production of subjects (e.g., priming and/or sentence completion). Furthermore, as noted in the introduction, few – if any - studies have examined children's use of inflections in such detail as the present study. Extending this elicited production method to other languages including those that have more complex present tense paradigms (e.g., Spanish with its different conjugation classes) would be informative.

Future research should also address the issue of how children's use of inflection becomes adult-like. The present study provides evidence that children move away from reliance on phonological analogy with neighbors and towards more abstract representations. It does not, however, address the issue of precisely how this change occurs, or what these more abstract representations look like. Do Finnish-speaking adults have, for example, an independently represented *[STEM]-n* construction or a cluster of exemplars tied together by functional as well as phonological similarity? In addition, even though adults seem to use these more abstract representations, there are probably circumstances in which they instead retrieve a ready-inflected form or apply phonological analogy to a close neighbor. Presumably, adults use a mixture of all three strategies, depending – among other factors – on the frequency of the target form, and hence the strength of its representation in memory. Further research is required to fully understand the complex relationship between these factors.

In conclusion, the present findings suggest that any successful account of the acquisition of verb morphology will need to include a role for rote-storage of

individual inflected forms as well as phonological analogy across such forms. Explaining how children move from this early stage characterized by rote-learning and errors in low frequency parts of the inflectional system to the fully-productive, error-free adult system remains a challenge for all theoretical approaches. It is to be hoped that future studies of inflectional morphology will cast more light on the relative balance of input-based learning and innate categories and formal rules; an issue that has important theoretical implications not only for accounts of morphological acquisition, but also for accounts of language acquisition in general.

**References**

Aguado-Orea, J. (2004). *The acquisition of morpho-syntax in Spanish: implications for current theories of development*. Unpublished doctoral dissertation, University of Nottingham, UK.

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*, 119–161.

Ambridge, B. (2010). Children's judgments of regular and irregular novel past tense forms: New data on the dual- versus single-route debate. *Developmental Psychology*, *46*, 1497-1504.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, England: Cambridge University Press.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*(4), 639-647.

Bates, E., & MacWhinney, B. (1989). Functionalism and the Competition Model. In B. MacWhinney, & E. Bates (Eds.), *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150-177.

Blom, W.B.T. & Wijnen, F.N.K. (2013). Optionality of finiteness: Evidence for a no overlap stage in Dutch child language. *First language*, *33*, 225-245.

Booij, G. (2010). Construction morphology. *Language and Linguistics Compass*, *4*(7), 543-555.

Bowerman, M. F. (1973). *Early Syntactic Development: A Crosslinguistic Study with Special Reference to Finnish*. London: Cambridge University Press.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge: Harvard University Press.

Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425-455.

Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J. L. (2013). Usage-based theory and exemplar representation. In Thomas Hoffman and Graeme Trousdale (eds.) *The Oxford Handbook of Construction Grammar*, pp. 49-69. Oxford University Press.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37, 575-596.

Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development*, *39*, 433-448.

Chomsky, N. (1995). *The Minimalist program*. *Cambridge, MA*: MIT Press.

Cipriani, P., Chilosi, A. M., Bottari, P. & Pfanner, L. (1991). *L'acquisizione della morfosintassi: Fasi e processi*. Padova: Unipress.

Clahsen, H., & Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences. In J. Meisel (Ed.), *The acquisition of verb placement. Functional categories and V2 phenomena in language acquisition* (pp. 181-224). Dordrecht: Kluwer.

Crain, S., & Thornton, R. (2012). Syntax acquisition. *WIREs Cognitive Science*, *3*, 185–203.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge University Press.

Dabrowska, E. (2008). The effects of frequency and neighbourhood density on adult speakers productivity with Polish case inflections: An empirical test of usage-

based approaches to morphology. *Journal of Memory and Language*, *58*, 931-951.

Dabrowska, E., & Szczerbinski, M. (2006). Polish children's productivity with case marking: the role of regularity, type frequency, and phonological diversity. *Journal of Child Language*, *33*, 559-597.

Deen, K. U. (2004). Productive agreement in Swahili: Against a piecemeal approach. In A. Brugos, M. R. Clark-Cotton & S. Ha (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development* (pp. 156-167). Boston: Cascadilla Press.

Gathercole, V. C. M., Sebastian, E. and Soto, P. (1999). The early acquisition of Spanish verbal morphology: across-the-board or piecemeal knowledge? *International Journal of Bilingualism*, 2-3, 133–82.

Guasti, M. T. (2004). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Hakulinen, A., Vilkuna, M., Korhonen, R., Koivisto, V., Heinonen, T., & Alho, I. (2004). *Iso suomen kielioppi.* Helsinki, Finland: Suomalaisen Kirjallisuuden Seura.

Harris, T. & Wexler, K. (1996). The optional infinitive stage in child English. In H. Clahsen (Ed.), *Generative perspectives on language acquisition* (pp. 1-42). Amsterdam: John Benjamins.

Hoekstra, T., & Hyams, N. (1998). Aspects of Root Infinitives. *Lingua, 106*, 81-112.

Janssen, D. P., Roelofs, A., & Levelt, W. J. (2002). Inflectional frames in language production. *Language and Cognitive Processes*, *17*(3), 209-236.

Karlsson, F., & Koskenniemi, K. (1985). A process model of morphology and lexicon. *Folia Linguistica*, *XIX(1*-2), 217-231.

Kirjavainen, M., Nikolaev, A., & Kidd, E. (2012). The effect of frequency and phonological neighbourhood density on the acquisition of past tense verbs by Finnish children. *Cognitive Linguistics*, *23*, 273-315.

Kunnari, S., Savinainen-Makkonen, T., Leonard, L. B., Mäkinen, L., Tolonen, A-K., Luotonen, M., & Leinonen, E. (2011). Children with Specific Language Impairment in Finnish: the use of tense and agreement inflections. *Journal of Child Language*, *38*, 999-1027.

Legate, J. A., & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, *14*, 315-344. doi:10.1080/10489220701471081

Leonard, L. B., Caselli, M. C., & Devescovi, A. (2002). Italian children's use of verb and noun morphology during the preschool years. *First Language*, *22*, 287-304.

Lukacs, A., Leonard, L. B., Kas, B., & Pleh, C. (2009). The use of tense and agreement by Hungarian-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 52*, 98-117.

Lust, B. (2006). *Child language: Acquisition and growth*. Cambridge: Cambridge University Press.

Maratsos, M. (2000). More overregularizations after all: New data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen & Xu. *Journal of Child Language*, *27*, 183-212.

Marchman V. A. (1997). Children's productivity in the English past tense: The role of

frequency, phonology, and neighborhood structures. *Cognitive Science*, *21*, 283-304.

Marchman, V. A., Wulfeck, B., & Weismer, S. E. (1999). Morphological productivity in children with normal language and SLI: a study of the English past tense. *Journal of Speech, Language and Hearing Research*, *42*, 206-219.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, *6*, 465-472.

Mielikäinen, A. (1984). Monikon 3.persoonon kongruenssi puhekielessä [Concord of 3rd person plural in present-day spoken Finnish]. *Virittäjä*, 162-175.

Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). Comparing different models of the development of the English verb category. *Linguistics, 36,* 807-830.

Pine, J. M., Conti-Ramsden, G., Joseph, K. L., Lieven, E. V. M. and Serratrice, L. (2008). Tense over time: testing the Agreement/Tense Omission Model as an account of the pattern of tense-marking provision in early child English. *Journal of Child Language*, 35, 55–75.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*, 456-463.

Pizzuto, E., & Caselli, M. C. (1992). The acquisition of Italian morphology: implications for models of language development. *Journal of Child Language*, *19*, 491-557.

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*, 1-56.

Räsänen, S. H. M., Ambridge, B., & Pine, J. M. (2014). Infinitives or bare stems? Are English-speaking children defaulting to the highest frequency form? *Journal of Child Language*, *41*, 756-779.

Rice, M. L. (2004). Growth models of developmental language disorders. In M. L. Rice & S. F. Warren (Eds.). *Developmental language disorders: From phenotypes to etiologies.* (pp. 207-240). Mahwah, NJ: Lawrence Erlbaum Associates.

Riionheimo, H. (2002). Havaintoja itäsuomalaisen lapsen morfologisista poikkeamista : verbintaivutus. *Virittäjä*, *106*, 259-270.

Rispoli, M., Hadley, P. A., & Holt, J. K. (2009). The growth of tense productivity. *Journal of Speech, Language and Hearing Research, 52*, 930-944.

Rubino, R. B., & Pine, J. M. (1998). Subject-verb agreement in Brazilian Portuguese: what low error rates hide. *Journal of Child Language*, *25*, 35-59.

Serra, M. & Solé, R. (1992). *Language acquisition in Spanish and Catalan children longitudinal study*. Universitat de Barcelona.

Skousen, R., Lonsdale, D. & Parkinson, D.B. (eds) (2002). *Analogical Modeling. An exemplar-based approach to language.* Amsterdam: John Benjamins.

Stavrakaki, S., & Clahsen, H. (2009). The perfective past tense in Greek child language. *Journal of Child Language, 36*, 113-142.

Theakston, A. L., Lieven E. V., & Tomasello M. (2003). The role of input in the acquisition of third-person singular verbs in English. *Journal of Speech, Language, & Hearing Research*, 46, 863-877. doi:10.1044/1092-4388(2003/067)

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*, 209-253.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, *106*, 23-79.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language*, *30*, 75-115.

**Table 1. Examples of probe items**

Examples of the probe items using the verb *katsoa* '*to look*'

| Verb inflection | Elicitation task | Expected response |
|---|---|---|
| 3sg pres | **Watching the video**<br>The experimenter asks: *Mitä koira tekee*?<br>'*What does the dog do*?] | *Koira katso/o*<br>*The dog looks* |
| 1sg pres | **Imitation of action**<br>The experimenter tells the child that now it is his/her turn to perform the action in the video. Whilst acting out the action, the Talking Dog asks: *Mitä sinä teet*?<br>[*What are you doing*?] | *(Minä) katso/n*<br>*(I) look* |
| 1pl pres | **Imitation of action**<br>The experimenter tells the child that now it is their turn to perform the action in the video together. Whilst acting out the action, the Talking Dog asks: *Mitä te teette*? [*What are you-pl doing*?] | *(Me) katso/mme*<br>*(We) look* |
| 2sg pres | **Watching the video**<br>The Talking Dog asks: *Mitä minä teen*?<br>[*What am I doing*?] | *(Sinä) katso/t*<br>*(You) look* |

| | | |
|---|---|---|
| 2pl pres | **Watching the video** | *(Te) katso/tte* |
| | The Talking Dog asks: *Mitä me teemme?* | *(You-pl) look* |
| | [*What are we doing*?] | |

**Table 2. Mean proportion of correct forms for each verb**

| | Glossary (low/high frequency status in brackets) | Number of children contributing data | Proportion correct inflection |
|---|---|---|---|
| *Aterioida* | *To eat (low)* | 7/41 | 0.86 |
| *Hymyillä* | *To smile (high)* | 35/41 | 0.84 |
| *Ilakoida* | *To be happy (low)* | 29/41 | 0.86 |
| *Iloita* | *To be happy (high)* | 31/41 | 0.83 |
| *Juoda* | *To drink (high)* | 38/41 | 0.92 |
| *Katsoa* | *To watch (high)* | 37/41 | 0.87 |
| *Kävellä* | *To walk (high)* | 39/41 | 0.85 |
| *Kisata* | *To play (low)* | 26/41 | 0.92 |
| *Kohota* | *To get up (low)* | 27/41 | 0.78 |
| *Kököttää* | *To stand (low)* | 28/41 | 0.85 |
| *Kulauttaa* | *To drink (low)* | 10/41 | 0.93 |
| *Lastata* | *To pack (low)* | 32/41 | 0.85 |
| *Leikata* | *To cut (high)* | 41/41 | 0.86 |
| *Liikuttaa* | *To take (low)* | 33/41 | 0.90 |
| *Löpistä* | *To talk (low)* | 27/41 | 0.83 |
| *Lukita* | *To lock (high)* | 34/41 | 0.85 |
| *Lyödä* | *To hit (high)* | 35/41 | 0.92 |
| *Maalata* | *To paint (high)* | 36/41 | 0.87 |
| *Myhäillä* | *To smile (low)* | 30/41 | 0.85 |

| | | | |
|---|---|---|---|
| *Nousta* | *To get up (high)* | 37/41 | 0.86 |
| *Nukkua* | *To sleep (high)* | 41/41 | 0.82 |
| *Pakata* | *To pack (high)* | 38/41 | 0.88 |
| *Panna* | *To put (high)* | 28/41 | 0.81 |
| *Pelata* | *To play (high)* | 36/41 | 0.82 |
| *Piiskata* | *To hit (low)* | 31/41 | 0.81 |
| *Puhua* | *To talk (high)* | 33/41 | 0.87 |
| *Saksia* | *To cut (low)* | 24/41 | 0.82 |
| *Salvata* | *To lock (low)* | 32/41 | 0.89 |
| *Seisoa* | *To stand (high)* | 39/41 | 0.85 |
| *Sijoittaa* | *To put (low)* | 30/41 | 0.87 |
| *Silmäillä* | *To watch (low)* | 30/41 | 0.77 |
| *Sivellä* | *To paint (low)* | 27/41 | 0.79 |
| *Syödä* | *To eat (high)* | 40/41 | 0.91 |
| *Talsia* | *To walk (low)* | 30/41 | 0.80 |
| *Uinua* | *To sleep (low)* | 15/41 | 0.85 |
| *Viedä* | *To take (high)* | 33/41 | 0.92 |

**Table 3. Error rates by inflectional target context**

| | 3sg | 1sg | 1pl | 2sg | 2pl |
|---|---|---|---|---|---|
| (a) Overall error rate | 0.46% | 10.34% | 11.67% | 14.38% | 35.83% |
| (b) Error rate excluding children who did not produce at least one instance of the target inflection | 0.46% | 9.71% | 11.68% | 13.96% | 31.84% |
| No. children contributing to (b) | 81 | 70 | 71 | 74 | 65 |

**Table 4. Errors broken down by inflectional target context**

| | Target inflection | | | | |
|---|---|---|---|---|---|
| Actual production | 3sg | 1sg | 1pl | 2sg | 2pl |
| 3sg | Correct | 6 | 25 | 57 | 68 |
| 1sg | 2 | Correct | 34 | *42* | 10 |
| 1pl | 0 | 23 | Correct | 0 | *38* |
| 1pl passive | 1 | 33 | Correct | 7 | *118* |
| 2sg | 1 | *38* | 13 | Correct | 113 |
| 2pl | 1 | 4 | *42* | 38 | Correct |
| 3pl | | | 1 | 1 | 1 |
| Total *N* errors | 5 | 104 | 115 | 145 | 348 |

**Table 5. Mixed-Effects Regression Models.**

*Model 1: Reduced model - Syllable length*

| Variable | β | *SE* | *z* | *p* |
|---|---|---|---|---|
| (Intercept) | 2.89 | 0.33 | 8.72 | < 0.001 |
| **Syllable length** | **-0.35** | **0.11** | **-3.22** | **0.001** |

*Note*. Model log likelihood = -1936. Random effects: Participant (*Var*=0.73, *SD*=0.85), Verb (*Var*=0.02, *SD*=0.14)

Bold values indicate that effect is statistically significant at $p < .05$ or greater.

*Model 2: Token frequency*

| Variable | β | *SE* | *z* | *p* |
|---|---|---|---|---|
| (Intercept) | 2.14 | 0.37 | 5.76 | < 0.001 |
| Syllable length | -0.15 | 0.12 | -1.25 | 0.211 |
| **Token** | **0.04** | **0.006** | **6.86** | **< 0.001** |

*Note*. Model log likelihood = -1884. Random effects: Participant (*Var*=0.81, *SD*=0.90), Verb (*Var*=0.04, *SD*=0.19)

Bold values indicate that effect is statistically significant at $p < .05$ or greater.

*Model 3: Token frequency + Morphophonological class size*

| Variable | β | *SE* | *z* | *p* |
|---|---|---|---|---|
| (Intercept) | 2.13 | 0.38 | 5.61 | < 0.001 |
| Syllable length | -0.15 | 0.13 | -1.10 | 0.271 |
| **Token frequency** | **0.04** | **0.006** | **6.86** | **< 0.001** |
| Morphophonological class size | -0.01 | 0.07 | -0.12 | 0.907 |

*Note*. Model log likelihood = -1884. Random effects: Participant (*Var*=0.81, *SD*=0.90), Verb (*Var*=0.04, *SD*=0.19)

Bold values indicate that effect is statistically significant at $p < .05$ or greater.

*Model 4: Token frequency + Morphophonological class size + Interactions*

| Variable | β | *SE* | *z* | *p* |
|---|---|---|---|---|
| (Intercept) | 1.84 | 0.37 | 5.01 | < 0.001 |
| Syllable length | -0.08 | 0.13 | -0.64 | 0.521 |
| **Token frequency** | **0.09** | **0.014** | **6.42** | **< 0.001** |
| Morphophonological class size | 0.06 | 0.07 | 0.94 | 0.348 |
| **Token frequency * Morphophonological class size** | **-0.03** | **0.007** | **-4.38** | **< 0.001** |

*Note*. Model log likelihood = -1873. Random effects: Participant (*Var*=0.83, *SD*=0.91), Verb (*Var*=0.02, *SD*=0.16)

Bold values indicate that effect is statistically significant at $p < .05$ or greater.

*Model 5: Token frequency + Morphophonological class size +Age + Interactions*

| Variable | β | *SE* | *z* | *p* |
|---|---|---|---|---|
| (Intercept) | -1.10 | 0.77 | -1.43 | 0.154 |
| **Age** | **6.69** | **1.57** | **4.27** | **< 0.001** |
| Syllable length | -0.10 | 0.13 | -0.78 | 0.436 |
| Token frequency | 0.10 | 0.08 | 1.27 | 0.205 |
| **Morphophonological class size** | **0.81** | **0.33** | **2.43** | **0.015** |

| | | | | |
|---|---|---|---|---|
| Token frequency * Morphophonological class size | -0.07 | 0.04 | -1.46 | 0.146 |
| Token frequency * Age | -0.003 | 0.18 | -0.02 | 0.986 |
| **Morphophonological class size * Age** | **-1.63** | **0.72** | **-2.28** | **0.023** |
| Token frequency *Morphophonological class size * Age | 0.6 | 0.09 | 0.63 | 0.528 |

*Note*. Model log likelihood = -1863. Random effects: Participant (*Var*=0.67, *SD*=0.82), Verb (*Var*=0.03, *SD*=0.16)

Bold values indicate that effect is statistically significant at $p < .05$ or greater.

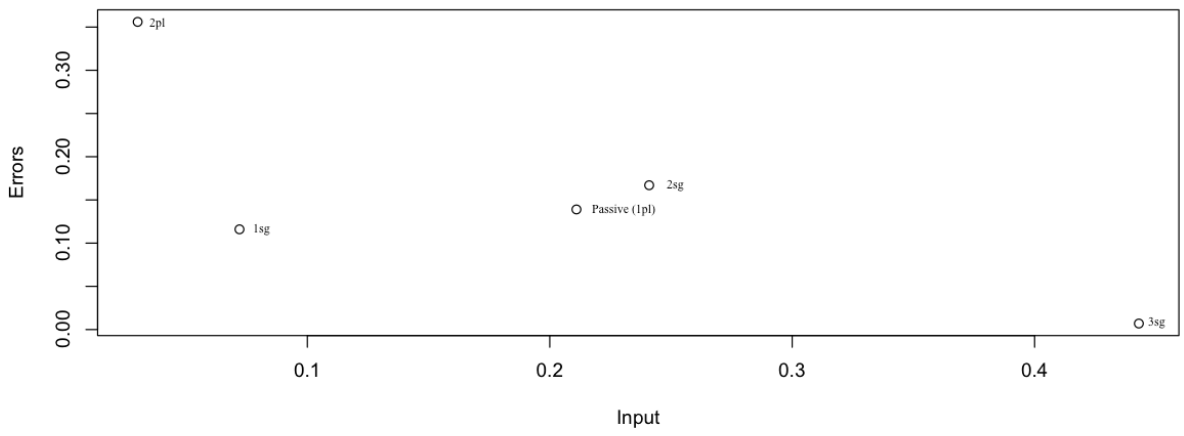

*Figure 1*. Illustration for the correlation between the adult input form frequency in a child-directed speech and children's error rate

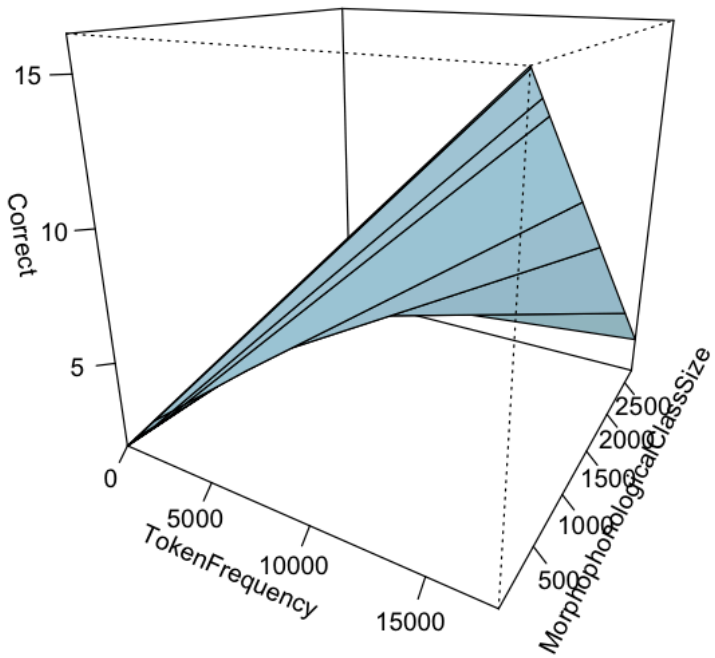


*Figure 2*. Illustration for the interaction between lexical token form frequency and morphophonological class size

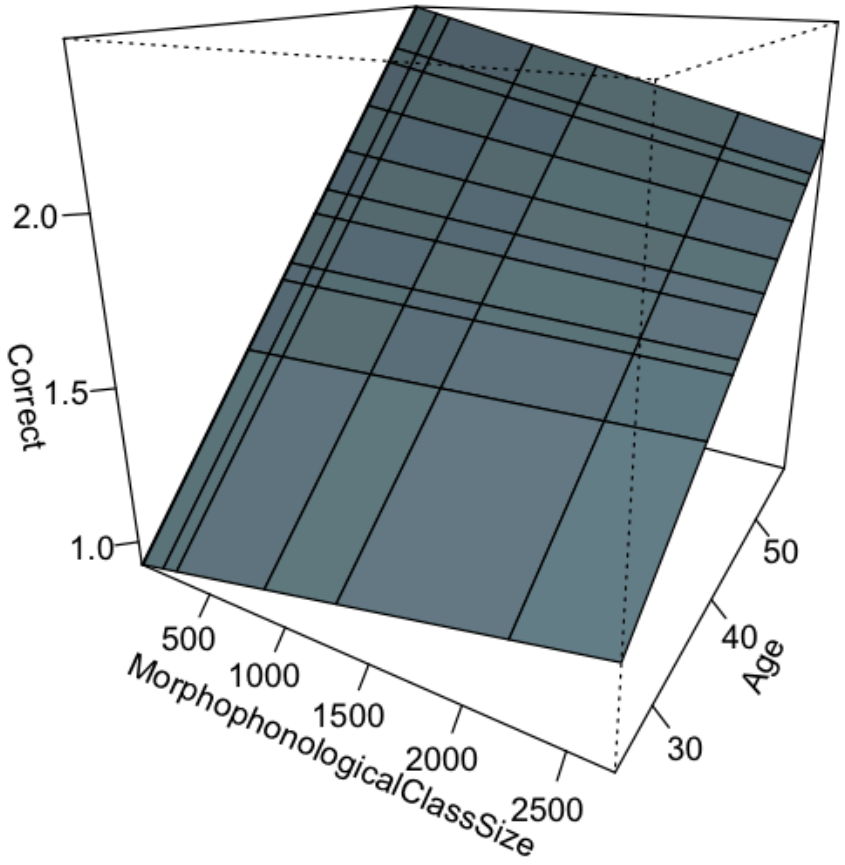


*Figure 3*. Illustration for the interaction between age and morphophonological class size

**Appendix A: Morphological verb classes**

I *Sano/a* verbs

*Sano/a* verbs, in which the infinitive inflection *–a* is placed after a short vowel, only have an infinitival stem (*sano/*). Thus, all the finite inflections are added directly to this infinitival stem. The stem may undergo consonant gradation (e.g., *antaa* 'to give' → *annan* 'I give'). This is the largest verb group with altogether 5754 verbs. However, it consists of eight subtypes, some of which have very low type frequencies (e.g., *kaartaa* verbs *N*=3)

II *Saa/da* verbs

Just like *Sano/a* verbs, *Saa/da* verbs have no separate inflectional stem, but finite inflections are added directly to the infinitival stem after removing the infinitival morpheme *–da*. The infinitival stem ends in a long vowel or a diphthong. This verb group can be considered as productive, as new polysyllabic verbs can come to this group. The total number of verbs in this group is 745, which consists of 15 highly frequent two-syllable verbs such as *voida* '*to be able'*, *syödä* '*to eat'* and *juoda* '*to drink'*. Two other highly frequent verbs in this group, *näh/dä* '*to see'* and *teh/dä* '*to do*' are however irregular as they have an inflectional stem that ends in *–ke* (e.g., *näkee* '*he sees'*) and are subject to consonant gradation. Altogether this verb type has three subtypes.

III *Nous/ta*, *Tul/la* and *Men/nä* verbs

*Nous/ta*, *Tul/la* and *Men/nä* verbs have an inflectional stem that ends in *–e*. This *–e* is added to the infinitival stem (e.g. *tul/la* '*to come'* → *tul/e/n* '*I come'*), and consonant gradation may occur depending on the stem (e.g., *ajatel/la* '*to think'* → *ajat**t**el/e/n* '*I think'*). Total number of verbs in the verb type is 1609, which consists of four subtypes.

IV *Huomat/a* verbs

*Huomat/a* verbs are a very productive class of verbs, as new verbs usually go to this group. These verbs are also known as contracted verbs because the complex relationship between the infinitival and the inflection stem. The final *–t* of the infinitive stem is changed to *–a/-ä* (e.g., *huomat/a* '*to notice'* → *huomaat* '*you notice*'), and as with the group III verbs, consonant gradation occurs only in the infinitival stem (e.g., *tavata* '*to meet'* → *tapaamme* '*we meet'*). The total number of verbs belonging to this verb type is 1067, consisting of three subtypes.

V *Tarvi/ta* verbs

*Tarvit/a* verbs resemble *Huomat/a* verbs in their infinitival form, but they differ in how their inflectional stem is formed. *Tarvit/a* verbs form their inflectional stem by adding *–se* to the infinitival stem (e.g., *tarvit/a* '*to need'* → *tarvit/se/mme* '*we need*'). This group of verbs is not very large (*N*=49), but it includes two frequent verbs, *tarvita* ('*to need*) and *häiritä* ('*to bother someone'*).

VI *Vanhet/a* verbs

*Vanhet/a* verbs are a rare group of verbs (*N*=143). These verbs form their inflectional stem by changing the final infinitival *–t* to *ne* (e.g., *vanhet/a* '*to get older'* → *vanhe/ne/vat* '*they get older'*). *Vanhet/a* verbs tend to be derived from adjectives, and have the meaning of becoming the adjective (e.g., *lämmet/ä* '*get warm'*; *vanhet/a* '*get older'*; *suuret/a* '*get bigger'*).

Table below summarizes the different morphophonological classes and their respective numbers

| **I *Sano/a* verbs** | **Number and example** | **Included in the study** | **II *Saa/da* verbs** | **Number and example** | **Included in the study** |
|---|---|---|---|---|---|
| *a* | *pamahtaa* 2714 | Yes | *One-syllable* | *voida* 15 | Yes |
| | *häätää* 318 | | *oi* | *tupakoida* 681 | Yes |
| | *ajaa* 51 | | *oi ~ oitse* | *luennoida* 49 | |
| | *yltaa* 7 | | | | |
| | *kaartaa* 3 | | | | |
| *e* | *hakea* 31 | | | | |
| *i* | *etsiä* 402 | Yes | | | |
| *o, u* | *puhua* 2228 | Yes | | | |
| Total | 5754 | | | 745 | |

| **III *Nous/ta, Tul/la and Men/nä* verbs** | **Number and example** | **Included in the study** | **IV *Huomat/a* verbs** | **Number and example** | **Included in the study** |
|---|---|---|---|---|---|
| *le* | *opiskella* 1329 | Yes | *aa* | *huomata* 885 | Yes |
| *se* | *nousta* 272 | Yes | *ua, oa, ea* | *katketa* 170 | Yes |
| *kse* | *syöstä* 3 | | *ia* | *hävitä* 12 | |
| *ne, re* | *purra* 5 | Yes | | | |
| Total | 1609 | | | 1067 | |
| **V *Tarvi/ta* verbs** | **Number and example** | **Included in the study** | **VI *Vanhet/a* verbs** | **Number and example** | **Included in the study** |
| *tse* | *tarvita* 49 | Yes | *ene* | *vanheta* 143 | |
| Total | 49 | | Total | 143 | |

**Appendix B: Verbs used in the present experiment and their characteristics**

| Verb | Glossary (low/high frequency status in brackets) | 3sg pres token frequency | 1sg pres token frequency | 1pl +passive pres token frequency | 2sg pres token frequency | 2pl pres token frequenc |
|---|---|---|---|---|---|---|
| *Aterioida* | *To eat (low)* | 41 | 0 | 8 | 0 | 0 |
| *Hymyillä* | *To smile (high)* | 1651 | 29 | 20 | 3 | 0 |
| *Ilakoida* | *To be happy (low)* | 119 | 0 | 3 | 0 | 0 |
| *Iloita* | *To be happy (high)* | 1570 | 66 | 116 | 2 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Juoda* | *To drink (high)* | 1060 | 121 | 609 | 23 | 11 |
| *Katsoa* | *To watch (high)* | 1105 | 1022 | 7409 | 148 | 37 |
| *Kävellä* | *To walk (high)* | 840 | 162 | 205 | 22 | 7 |
| *Kisata* | *To play (low)* | 716 | 352 | 153 | 0 | 0 |
| *Kohota* | *To get up (low)* | 1611 | 2 | 5 | 0 | 0 |
| *Kököttää* | *To stand (low)* | 39 | 0 | 6 | 2 | 0 |
| *Kulauttaa* | *To drink (low)* | 21 | 0 | 4 | 0 | 0 |
| *Lastata* | *To pack (low)* | 50 | 0 | 144 | 2 | 0 |
| *Leikata* | *To cut (high)* | 1919 | 30 | 1904 | 16 | 5 |
| *Liikuttaa* | *To take (low)* | 358 | 0 | 23 | 0 | 0 |
| *Löpistä* | *To talk (low)* | 1 | 0 | 0 | 0 | 0 |
| *Lukita* | *To lock (high)* | 33 | 0 | 40 | 2 | 0 |
| *Lyödä* | *To hit (high)* | 2343 | 71 | 917 | 23 | 5 |
| *Maalata* | *To paint (high)* | 843 | 98 | 327 | 10 | 0 |
| *Myhäillä* | *To smile (low)* | 1079 | 0 | 3 | 0 | 0 |
| *Nousta* | *To get up (high)* | 17755 | 95 | 429 | 19 | 2 |
| *Nukkua* | *To sleep (high)* | 613 | 182 | 137 | 23 | 4 |
| *Pakata* | *To pack (high)* | 373 | 7 | 345 | 6 | 0 |
| *Panna* | *To put (high)* | 2298 | 134 | 3110 | 22 | 12 |
| *Pelata* | *To play (high)* | 7839 | 414 | 5216 | 84 | 10 |
| *Piiskata* | *To hit (low)* | 125 | 17 | 19 | 0 | 0 |
| *Puhua* | *To talk (high)* | 7686 | 429 | 7977 | 78 | 39 |
| *Saksia* | *To cut (low)* | 3 | 0 | 6 | 0 | 0 |
| *Salvata* | *To lock (low)* | 28 | 0 | 0 | 0 | 0 |
| *Seisoa* | *To stand (high)* | 2290 | 88 | 103 | 76 | 8 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Sijoittaa* | *To put (low)* | 4594 | 32 | 2350 | 12 | 0 |
| *Silmäillä* | *To watch (low)* | 103 | 12 | 10 | 0 | 0 |
| *Sivellä* | *To paint (low)* | 39 | 4 | 67 | 0 | 0 |
| *Syödä* | *To eat (high)* | 2957 | 219 | 1406 | 93 | 13 |
| *Talsia* | *To walk (low)* | 13 | 2 | 2 | 0 | 0 |
| *Uinua* | *To sleep (low)* | 64 | 0 | 0 | 12 | 0 |
| *Viedä* | *To take (high)* | 16303 | 288 | 5402 | 65 | 18 |

| Verb | Morphophonological class size | Adult frequency rating (1= Not frequent at all; 10 = Very frequent) | Syllables in infinitive |
|---|---|---|---|
| *Aterioida* | 681 | 3.57 | 5 |
| *Hymyillä* | 1329 | 8.98 | 3 |
| *Ilakoida* | 681 | 2.69 | 4 |
| *Iloita* | 49 | 6.92 | 3 |
| *Juoda* | 15 | 9.71 | 2 |
| *Katsoa* | 2228 | 9.71 | 3 |
| *Kävellä* | 1329 | 9.71 | 3 |
| *Kisata* | 885 | 5.08 | 3 |
| *Kohota* | 170 | 5.43 | 3 |
| *Kököttää* | 2714 | 4.61 | 3 |
| *Kulauttaa* | 2714 | 4.12 | 3 |
| *Lastata* | 885 | 6.10 | 3 |
| *Leikata* | 885 | 8.76 | 3 |

| | | | |
|---|---|---|---|
| *Liikuttaa* | 2714 | 7.67 | 3 |
| *Löpistä* | 272 | 3.98 | 3 |
| *Lukita* | 49 | 8.14 | 3 |
| *Lyödä* | 15 | 8.69 | 2 |
| *Maalata* | 885 | 8.33 | 3 |
| *Myhäillä* | 1329 | 5.29 | 3 |
| *Nousta* | 272 | 8.80 | 2 |
| *Nukkua* | 2228 | 9.76 | 3 |
| *Pakata* | 885 | 8.73 | 3 |
| *Panna* | 5 | 5.00 | 2 |
| *Pelata* | 885 | 9.19 | 3 |
| *Piiskata* | 885 | 5.85 | 3 |
| *Puhua* | 2228 | 9.52 | 3 |
| *Saksia* | 402 | 3.55 | 3 |
| *Salvata* | 885 | 1.85 | 3 |
| *Seisoa* | 2228 | 9.44 | 3 |
| *Sijoittaa* | 2714 | 7.17 | 3 |
| *Silmäillä* | 1329 | 6.15 | 3 |
| *Sivellä* | 1329 | 5.21 | 3 |
| *Syödä* | 15 | 9.65 | 2 |
| *Talsia* | 402 | 3.89 | 3 |
| *Uinua* | 2228 | 3.54 | 3 |
| *Viedä* | 15 | 9.13 | 2 |

## Footnotes

[i] Note that the person/number context might be indicated by the presence of a subject (e.g., *I*....); but it might not. Many languages (including Finnish, Italian, Spanish and Catalan) allow speakers to drop subjects when they can be easily inferred from the discourse. This subject-drop does not absolve the speaker of her responsibility to provide an appropriate person/number marked verb form, even though there is no "agreement" with an overt subject. Thus, all of the generativist accounts and studies discussed in this section include as instances of correct "agreement" utterances in which the intended person/number context is inferred from the surrounding discourse with no overt subject present. Consequently, we follow this standard practice in the present study.

[ii] It should be noted that when the subject is not nominative, the 3sg form must be used. This applies for instance to possessive and necessive constructions, which are frequent in the input.

[iii] With regard to vowel harmony, front vowels (/*ä ö y*/) cannot co-occur with back vowels (/*a o u*/). For example, the verb *syö/dä* '*to eat*' has *–dä* as an infinitival ending because the word stem contains front vowels, whereas *juo/da* '*to drink*' has *–da* as an infinitival ending because of the back vowels in the stem of the verb. Consonant gradation refers to deletion and lenition of consonants when the verbs are inflected. This phenomenon occurs when long voiceless stops *pp*, *tt* and *kk* are shortened to *p*, *t* and *k*, respectively. This is known as *quantitative* gradation. In contrast, short voiceless stops *p*, *t* and *k* are weakened in several *qualitative* ways (e.g., *p* → *v*; *p* → *m*; *t* → *d*). The conditions for consonant gradation are rule-governed, but very complex.

[iv] Although such forms were not elicited in the present study, there is an increasing tendency in spoken Finnish to replace the infinitival form of *Huomat/a* verbs (see Appendix A) with the 3sg form in compound finite structures such as *En jaksa pakkaa* (for *pakat/a*) '*I can't be bothered to pack* and *Aloitan pakkaa* (for *pakkaamaan*) '*I start to pack'*.

[v] The proportion of 3pl forms in 17 transcriptions of child-directed speech is 0.30% (total number of

present tense forms = 1748).

[vi] Interestingly, the rate of stem errors also increased with age ($\beta$ = 6.84, $SE$ = 1.98, $z$ = 3.46, $p$ < .001). However, this finding is in line with what is known about the development of the Finnish inflectional system: as children's speech develops, they often overgeneralize the phonological alternations such as consonant gradation incorrectly, especially with passives (e.g., Riionheimo, 2002). In the present study too, an analysis of the stem errors revealed that such errors were significantly more frequent in 1pl passive contexts than in any other contexts ($p$ < .001), with no other differences between inflectional contexts observed.