Core morphology in child directed speech

Crosslinguistic corpus analyses of noun plurals*

Dorit Ravid, Wolfgang U. Dressler, Bracha Nir-Sagiv,
Katharina Korecky-Kröll, Agnita Souman, Katja Rehfeldt,
Sabine Laaha, Johannes Bertl, Hans Basbøll and Steven Gillis

1. Introduction

Learning inflectional systems is a crucial task taken up early on by toddlers. From a *distributional* point of view, inflection is characterized by high token frequency, and general and obligatory applicability (Bybee 1985). From a *semantic* point of view, inflection exhibits transparency, regularity and predictability. These aspects of inflection render it highly salient for young children and facilitate the initial mapping of meaning or function onto inflectional segments. At the same time, many inflectional systems are also fraught with *morphological and morpho-phonological* complexity, opacity, inconsistency, irregularity, and unpredictability. These *structural* aspects of inflection constitute a serious challenge to the successful launching of this central function of human language.

Most studies of inflectional morphology start from an analysis of the adult system, and reason from that system the *when* and *how* of children's acquisition. However, the discrepancy between the complexity of the mature system, on the one hand, and the need to facilitate acquisition, on the other, needs to be resolved. Child Directed Speech (CDS) – simply defined as input to children from caregivers and early peer-group – has been shown to account for emerging lexical and morpho-syntactic features in

* For German and Hebrew: An important part of this work has been funded by the mainly experimental project Nr. P17276 "Noun development in a cross-linguistic perspective" of the Austrian Science Fund (FWF). Invited by Heike Behrens to contribute to this volume on the importance of the input children receive, we limited ourselves to longitudinal data only. For Dutch: Preparation of this paper was supported by a grant from the FWO (Flemish Science Foundation), contract G.0216.05. For Danish: Part of the Danish work was funded by the Carlsberg Foundation.

child language (Gallaway and Richards 1994; Ninio 1992; Ziesler and Demuth 1995).[1] The literature indicates that such linguistic input to young children consistently differs from speech among adults (Cameron-Faulkner, Lieven and Tomasello 2003; Gleitman, Gleitman, Landau and Wanner 1988; Morgan 1986; Snow 1995): it presents children with those aspects of the system which are particularly frequent, transparent, regular and consistent. These could make the child's job of understanding what the system is about and how it works much simpler.

We term these aspects of the adult inflectional system that are most easily transmitted to children *core morphology*. In the current study we consider core morphology within the domain of plural inflection in nouns. Specifically, we will show that across the languages we investigate here, the way the system is represented in CDS provides the child with clear and consistent information regarding its distributional aspects. This refers to the conditions for the distribution of types of plural suffixes as well as to the token-frequency of unproductive plural patterns. To the best of our knowledge, no cross-linguistic work has to date been carried out to document, define and analyze the nature and distribution of core morphology in Child Directed Speech and / or in young children's output. In our view, such work requires a systematic longitudinal analysis of spontaneous speech data of the type presented here: a crosslinguistic comparison of noun plurals in the input to, and output of, young children learning German, Dutch, Danish, and Hebrew.

Our concept of *core morphology* is clearly different in nature, scope and function from Chomsky's (1980) notion of *core grammar* (Joseph 1992), which equals innate Universal Grammar (also called the *Narrow Language Faculty* – Chomsky 1995; Fitch, Hauser and Chomsky 2005). Core grammar is language-specific only insofar as universally open parameter values are fixed in one of the universally given options. While both core morphology and core grammar relate to acquisition and psycholinguistic modelling in general, we do not share Chomsky's concepts of luxurious grammatical innateness, of the logical problem of learnability, or of insufficient and erroneous input evidence (MacWhinney 2004).

An older concept, only partially comparable to ours, is the Prague School notion of the *centre* of a linguistic system, as opposed to its periphery (Daneš 1966; Popela 1966). The overlapping criteria for the appurtenance of a morphological construction to the centre of a language are its prototypicality, its high degree of integration into a (sub)system (cf. the notion of system adequacy in Natural Morphology (Kilani-Schoch and Dressler 2005), its high type and token frequency and productivity – understood as applicability of a pattern to any new word that fits the structural description of the

1. In a recent, pertinent discussion on Info-Childes (4.12.2006), Dan Slobin commented that he preferred the term "exposure language" to other terms such as "input" (which assumes the child takes everything in), "motherese" and "caregiver talk" (which exclude talk from non-parents and non-caregivers), and "child directed speech" (which excludes what children learn from overheard speech). However, given later commentaries on CDS as a register, he conceded that this is a compact and convenient term. All participants commented on the need to specify the linguistic characteristics of CDS.

pattern (or of the input of a morphological rule). In the later literature, productive patterns were regarded as the core of morphology (and the rest of the grammar) by Dressler (1989; 2003) and Bertinetto (2003 :191ff), that is, unproductive patterns were regarded as marginal, inactive lexically stored parts of grammar.

Age of acquisition plays a crucial role in our current conception of *core morphology*. As pioneered by Jakobson (1941) and empirically investigated in abundant psycholinguistic research, early-emerging linguistic patterns are better stored and faster accessed by adults than what is acquired later on (Bonin, Barry, Méot and Chalard 2004; Burani, Barca and Arduino 2001; Lewis, Gerhard and Ellis 2001; Zevin and Seidenberg 2002). Early acquired patterns evidently depend on more limited input than later acquisition, in two senses: Firstly, the *amount* of tokens instantiating a morphological category or system is smaller than their number in Adult Directed Speech and speech addressed to older children; and secondly, their *variety* – that is, their different types and subtypes within and across categories – focuses on the most prototypical members of the category.[2]

1.1 Noun plurals in acquisition

Our window onto core morphology in this chapter is the path leading to the acquisition of *noun plurals* in three Germanic languages – Austrian German, Danish and Dutch – and one Semitic language, Hebrew. Plural formation is a basic category that emerges and develops early on in child language (Berman 1981; Ravid 1995; Stephany 2002). It has a large cross-linguistic distribution, including sign languages (Pfau and Steinbach 2006) and often exhibits much structural complexity (Corbett 2000). It plays a central role in the morphology of noun phrases and as the trigger of grammatical agreement. Plurals are signaled on nouns as the heads of noun phrases, if nouns carry any morphological marking in the respective language. Plural marking is the most basic morphological marker on nouns: if a language has a single category of morphological marking on the noun, it is grammatical number. Since singular marking is often zero, with duals having a much smaller distribution, plural is the central number marking in the world's languages. Accordingly, plural emerges as one of the earliest categories in child language development (Brown 1973; Slobin 1985c), and the path to its acquisition has been the topic of many studies and much controversy (Clahsen, Rothweiler, Woest and Marcus 1992; Marcus, Brinkmann, Clahsen, Wiese and Pinker 1995; Marcus, Pinker, Ullman, Hollander, Rosen and Xu 1992). The main concern in the current study is how children faced with complex and often inconsistent systems are able to 'break into the system' at the earliest stages of morphological acquisition.

2. By prototypicality we mean here relatively high type frequency and/or token frequency, i.e. a medium amount of token frequency is necessary for allowing high type frequency to establish a prototype, but if there is only low type frequency, then high token frequency overrules it and establishes by itself a prototype.

1.1.1 *Dual-route accounts*

For the acquisition and representation of English plurals, it is relatively easy to argue for the adequacy of a dual-route model account to explain how plurals are acquired and represented. This view, as proposed by Pinker (1999), assumes that regular forms are computed in the grammar by combinatorial operations that assemble morphemes and simplex words into complex words and larger syntactic units (Clahsen 1999; Marcus 2000; Sahin, Pinker and Halgren 2006). An important feature of this view is the dissociation of singular stem (base) and suffix as distinct symbolic variables (Berent, Pinker and Shimron 2002; Pinker and Ullman 2002). Regular plurals are thus productively generated by a general operation of unification, concatenating plural *-s* with the symbol N and inflecting any word categorized as a noun.

Under this view, irregular forms behave like words in the lexicon, that is, they are acquired and stored like other words with the plural grammatical feature incorporated into their lexical entries. Learning irregular forms is governed by associative memory, which facilitates the acquisition of similar items and superimposes the properties of old items on new ones resembling them. A stored inflected form blocks the application of the rule to that form, but elsewhere the rule applies to any item appropriately marked. At some point in acquisition English-speaking children would extract from the input generalizations for the formation of the sibilant plurals, the only productive and *default* pattern. Plural minor patterns and exceptions are truly infrequent in English as both types and tokens: the very few cases of umlaut (e.g. *foot – feet*, *mouse – mice*) and *-en* plurals (*child – children*) relevant to children would be rote-learned and remain separately stored words with the feature [plural] incorporated into their lexical entries.

1.1.2 *Challenges to the dual-route*

Unfortunately, this dual-route account cannot be easily extended to accommodate all of the four languages analyzed in this contribution (nor to the noun and verb inflection systems of, say, Slavic languages). For example, the attribution of a dual-route model to *German* (notably by Bartke, Marcus and Clahsen 1995; Clahsen 1999) assumes *-s* plurals to be the default, rule-derived form. However, these studies have not come to grips with the fact that across the literature on German-learning children, and for all Austrian ones described so far, *-s* plurals are neither the first ones to emerge, nor are they the only ones to be overgeneralized. Acquiring German plurals is better accounted for by single-route models (including schema-based models), which are also compatible with a gradual continuum between fully productive and unproductive plurals (Laaha, Ravid, Korecky-Kröll, Laaha and Dressler 2006).

Dutch plurals are difficult (if not impossible) to account for in a dual-route model. First of all, the Dutch plural is incompatible with a single default, since it has two suffixes (*-en* and *-s*), which are considered to be in complementary distribution (Baayen, Schreuder, De Jong, and Krott 2002; Booij 2001; De Haas and Trommelen 1993; van Wijk 2002; Zonneveld 2004; but see Bauer 2003). The distribution of the two suffixes is determined by the phonological structure of the singular, and more specifically, by

the word-final segment as well as the word's stress pattern. In other words, a noun's regular plural suffix is determined on the basis of its phonological profile. Thus, both suffixes are productive in their respective phonological domain, which makes them both candidates for default application. Linguistic analysis reveals that, besides productivity, both suffixes have the characteristics of a default inflectional pattern (Baayen, Dijkstra and Schreuder 1997; Baayen *et al.* 2002; Zonneveld 2004).

Even staunch advocates of the dual-route model observe that there is no single default in this case: Pinker and Prince (1994) remark that "the two affixes have separate domains of productivity... but within those domains they are both demonstrably productive" and call it "an unsolved but tantalizing problem." Pinker (1999) writes: "Remarkably, Dutch has two plurals that pass our stringent tests for regularity, *-s* and *–en*... Within their fiefdoms each applies as the default." Thus, Dutch plurals appear to deviate from the dual-route account in at least two respects: (1) there are two defaults instead of one; and (2) plural formation cannot be seen as the 'blind' application of a symbolic rule to the category N, since phonological information is needed in order to decide on the choice of the affix (similar to what is well-known for inflection in Slavic languages). The latter is not an enigma: recently, Keuleers, Sandra, Daelemans, Gillis, Durieux and Martens (2007) have shown that Dutch-speaking adults also use orthographic information in order to decide about which suffix to use.

Finally, *Hebrew* plurals too pose a challenge to the dual-route model, from a different perspective. Two studies test and analyze plural formation in a small number of Hebrew noun categories (Berent, Pinker and Shimron 1999, 2002). The authors regard suffix regularity and base change as independent of each other, concluding that they represent two different mental computations: symbolic operations versus memorized idiosyncrasies. The problem is, that the Berent *et al.*'s analysis hinges on viewing the base- and stress-preserving, masculine plural as the default Hebrew plural – an assumption tested, as in German and English, on proper names homophonous with common nouns. Pluralization of proper names (e.g., *Dov)* would yield a form extremely 'faithful' to the singular base – no base change, no stress shift – with the masculine *-im* suffix. This is supposed to constitute the default Hebrew plural. Under the assumption that defaults constitute part of the plural system of a language, this test both overshoots and falls short of actually accounting for Hebrew plural formation (Ravid 2006), since it yields a non-Hebrew form. A critical factor is the fact that native Hebrew plurals – like all linear nominal suffixes[3] – always shift stress to the final syllable (e.g., *dov – dubím* 'bears'). Suffixation that fails to obey stress shift cannot be regarded as part of native Hebrew morphology, not to mention being considered a default plural. Moreover, the sensitivity of Hebrew suffix type to base-final phonology would lead to completely

3. Failure to move stress to the final syllable ("preserve stem faithfulness") in non-native words is not plural-specific and is a general feature of Hebrew nominal morphology: Compare foreign-based denominal adjectives *normáli* 'normal' or *fatáli* 'fatal' with native ultimate stressed *tsiburi* 'public'.

un-Hebrew forms under the proper name test. Thus for example *–it* final proper names such as *Maskít* would completely preserve base form and take masculine *-im* to yield *Maskítim* instead of undergoing *t*-deletion and stress shift and taking feminine *–ot* to yield *maskiyót* (Ravid 1995). *Maskítim* constitutes a plural form completely incompatible with native Hebrew morphology beyond toddlerhood (Berman 1985; Levy 1980). In general, plural formation of proper nouns is marginal both in plural use and in regard to morphological grammar in general. Thus, what is a default in plural formation (and inflection in general) should not be judged by what occurs in proper names.

Against this background, we now examine how single-route models handle plural formation (e.g., Daugherty and Seidenberg 1994; Plunkett and Marchman 1991; Rumelhart and McClelland 1986). Under this view, the learning network improves performance over many learning trials, resulting in a gradual developmental process where overgeneralization is conditioned by linguistic experience coupled with the similarity of the exemplar being learned to others already stored, its consistency and salience, as well as by frequency. Such single-route mechanisms can predict how grammatical representations are acquired. This cannot be said for dual-route models, which assume that children (like adults) eventually use a default rule and an associative memory system – but do not explain which mechanism accounts for how the default rule is acquired. Given these varied challenges to the dual route model, we adopt a single-route approach to plural acquisition.

We now turn to the problem of complexity in the plural systems under investigation, in order to assess the challenges faced by young learners.

1.2 Complexity in the formation of noun plurals

Plural formation takes on different degrees of complexity in the world's languages. For example, Turkish plural formation is most simple and homogeneous, involving just one, biunique suffix and almost no change in the nominal base; concomitantly plural emerges and consolidates early on in Turkish (Stephany 2002, with references). English plural formation is also relatively morphologically homogeneous, insofar as sibilant plurals represent the clear default and the only productive plural formation type with overwhelming type frequency. The three allomorphs in English *(-z, -s, Iz)* can be accounted for in a purely phonological way. However, plural formation of many other languages, including those represented in the current study, is much more complex, but to date, no overall measures of classifying degree of complexity have been proposed.

Two important facets of plural systems which contribute to their complexity and which children eventually have to learn are (1) plural suffix application and (2) subsequent changes to the base. For example, Hebrew singular masculine *iš* 'man' takes the plural suffix *-im*, and consequently changes the base to *anaš*, yielding plural *anaš-ím*. However, the scope of this chapter restricts us to focusing on plural suffix application in acquisition. This chapter thus presents a method of assessing complexity of *plural suffixation* in the four languages under investigation, to be used in the analyses of CDS and children's output.

Our comparative framework starts from the assumption that two recurrent factors are the most important ones for predicting the application of suffixation in our languages: *sonority* and *gender*. Phonological conditions have always been considered important for predicting suffixation patterns in many languages, but often not in any way that respects phonology systematically (a notable exception is palatality in Slavic languages). We propose the *sonority scale* (Goldsmith 1995) as one organizing phonological principle playing an important morphological role in all of the languages of this study. The sonority scale is a predictor of the order of segments within the syllable: the prototypical peak, i.e. the centre of the syllable, is (phonetically) a vowel, and among the consonants, obstruents (with noise, such as /p/ or /s/) are furthest away from the centre, whereas sonorants (noise-free, such as /l/, /m/) are closer to the centre. Our tables with sonority illustrate where on the sonority slope (from the peak rightwards) the final segment of the base is situated. This mirror-image of sonority in the syllable, with a peak in the middle and slopes to each side, is combined with inherent sonority (which does not predict order of segments in the syllable): stressed, low and full vowels are inherently more sonorous than unstressed, high and reduced vowels, respectively. Only the distinct position of Hebrew /t/ and /n/ cannot be derived from the sonority scale.

A second factor, shared by three of our four languages (German, Danish and Hebrew) is *gender* of the singular noun, a factor well-known for many Indo-European languages but often underrated for Germanic languages (Harbert 2006 :93, 96), with the exception of German (Köpcke 1993; Wegener 1999). We restrict our current analysis to these two factors since they allow us to put the four languages into the same perspective.

To illustrate how gender and degree of sonority of the base-final phoneme interact in determining the application of suffixation, Table 1 presents a fragment of German, consisting of four possible intersections of gender and sonority:

Table 1. A fragment of the interaction between gender and sonority in Austrian German

Sonority / **Gender**	**Obstruents**	**Schwa**
Feminine	Subregular: *-(e)n*, *-s*	Regular: *-n*
	Irregular: *-e*	Irregular: ø
Masculine	Subregular: *-e*, *-(e)n*, -s	Subregular: ø, *-n*

The four cells in Table 1 present the notion of regularity of suffixation as defined in the present context: the conditions under which rules (as formal expression of inflectional patterns) apply. Thus, the degree of regularity of suffixation is in fact the degree of predictability of the application of a specific suffixation rule in a given cell resulting from the interaction of sonority and gender (cf. Monaghan and Christiansen, *this volume,* for further discussion of multiple cue integration). If there is a clear default for

one productive suffixation to apply, we have regularity. For example, consider the suffixation of *-n* after feminine nouns ending in schwa in Table 1, as in *Orange-n* 'orange-s'. If any other rule applies in the same sonority-gender cell, we have irregularity, for example, feminine nouns ending in schwa with a zero suffix (e.g. *Mütter* 'mother-s'). But if two or more suffixation rules apply productively in the same cell (applying either optionally or alternatively to the same words or in complementary lexical distribution) we have subregularity. Thus both plural *-e* and *-s* may apply to the masculine noun *Park,* Pl. *Park-e, Park-s* 'park-s', and in other words *–en*, as in *Prinz-en* 'prince-s'.

Thus, based on Laaha *et al.* (2006 :280), we first distinguish between plural suffixations which freely apply, under a specific combination of gender and word-final phonology, to new words and are thus *productive*, and those which do not, and are thus *unproductive* – which we classify as *irregular*. Second, we distinguish between cells where just one productive plural suffixation pattern occurs (irrespective of whether there are some irregular exceptions) and those where two (or more) productive patterns compete. In the first case, we have a regular pattern (which is *fully predictable*, with possible irregular exceptions which have to be memorized according to all linguistic and psycholinguistic models); in the second case we identify two (or more) *subregular* patterns whose selection is unpredictable.

Our approach to the puzzle of noun plural learning thus starts out from this rich and complex view of gender x sonority in mature systems as the target of children's acquisition in the four study languages. The aim of this chapter is to establish empirically in what way exactly core morphology facilitates acquisition by identifying the domain of core morphology within mature noun plurals systems; that is, to determine to what extent and in what ways plural input to young children is restricted.

2. Language systems

This section describes the application of plural suffixation as a function of gender and sonority in the four languages under investigation. While the general scale of base-final sonority guides us across the board in the four languages, the actual set of categories and segments manifesting the sonority scale and appearing in the top row of Tables 2–5 below are each dictated by plural formation in the specific language under consideration. In the same way, gender, the other axis creating the grid for plural formation (if the language has it), is also presented from a language-specific perspective.

The analysis of the Danish language system is original in its account for morphology departing exclusively from sound structure, and not via the written language, and in its use of base-final sonority (systematically) and in the application of our common gender and base-final sonority framework. The analysis of the German plural system is new in its classification of *regular, subregular* and *irregular* suffixations, in its extension of phonological conditioning from word-final vowels to consonants, and in the introduction of the sonority hierarchy. The analysis of the Hebrew system is completely

new in the distinction it makes between regular and irregular plural suffixation, on the one hand, and gender-specific subregular patterning, as well as in the application of the sonority hierarchy to Hebrew plurals. The analysis of plural formation in Dutch provided here is fully in agreement with the linguistic descriptive tradition, in which two factors are considered to determine the choice of the plural suffix, viz. the final segment of the singular and the word's rhythm. This analysis dates back to Van Haeringen (1947), and since then analyses of plural formation have always stressed the importance of these two factors to different degrees (see De Haas and Trommelen 1993; Haeseryn, Romijn, Geerts, de Rooij and van den Toorn 1997 among others). Recently Van Wijk (2002) analyzed a corpus of written Dutch in order to establish where the balance lies between the rhythmic and the segmental factors.

2.1 Dutch plural formation

Plural formation of Dutch nouns consists in adding a suffix to the singular. There are two productive suffixes: *-en* /ə(n)/ and *–s* /s/, which are (largely) in complementary distribution.[4] Table 2 shows the distribution of the plural suffix according to the sonority scale only, since gender does not play a role in plural formation in Dutch. However there is an interesting interplay between the final segment(s) and the stress pattern of the word, and hence, for most types of words there is only subregularity (De Haas and Trommelen 1993; Van Wijk 2002).

Table 2. Sonority in Dutch

Obstruent	Sonorant	Schwa	Full Vowel
Subregular: *-en, -s*	Subregular: *-en, -s*	Regular: *-s*	Subregular: *-en, -s*
Irregular: *-en, -s*	Irregular: *-en, -s*	Irregular: *-en*	Irregular: *-en, -s*

Words ending in an obstruent take *–en* as their plural suffix if stress is on the final syllable, and *–s* if stress is on a pre-final syllable, so that the resulting plural form is a trochee. Thus, these patterns define the subregularity. But as Van Wijk (2002) points out in her corpus study: neither subregularity is exceptionless, which entails that both suffixes are also irregular. That is, *-s* is irregular for words with final stress and *–en* for words with prefinal stress.

Words ending in a sonorant tend to take the *–en* suffix when preceded by a full vowel and *–s* when preceded by a schwa. The latter regularity is very strong, though some of these words can take both suffixes (without an apparent meaning difference),

4. A third suffix, viz. *-eren*, is not productive anymore and only 12 nouns are pluralized with *-eren*. In addition, there are non-germanic plural markers as in *collega – collegae* ('colleague'), *musicus – musici* ('musician'). These are all not productive and are often replaced by a plural *–s/-en*: *collega – collegae – collegas*.

such as *appel – appel-s / appel-en* 'apple-s'. Thus, words ending in a schwa show a very straightforward picture: they take *–s* as a rule, though quite a few of these words can take the *–(e)n* plural as well: *syllabe – syllabe-s – syllabe-n* 'syllable-s'. The former has many exceptions, some of which can be explained by the metrical regularity that plurals are expected to end in a trochee, but still others are plain exceptions: *oom – oom-s* 'uncle-s', *roman – roman-s* 'novel-s'. Finally, diphthong-final words predominantly prefer the *–en* suffix (irrespective of the stress pattern of the word, e.g., *aardbei – aardbei-en* 'strawberrie-s' [artbEi], *bij – bij-en* 'bee-s' [bEi]), while words ending in a full vowel take *–s* (e.g., *positie – positie-s* 'position-s' [pozisis]). Again there are many exceptions, such as *zee – zee-en* 'sea-s', *koe* [ku:] *– koe-en* 'cow-s'.

2.2 German plural formation

The system of noun pluralization in German consists of more phonologically unrelated plural allomorphs than Dutch, also with no single clearly dominant form. German noun plurals are formed by the four different suffixes *-s, -(e)n, -e, -er* or by zero. The three latter ones may combine with umlaut (base vowel change), disregarded here since this chapter is not concerned with base changes.

Table 3. Interaction of gender and sonority in Austrian German[5]

Sonority / Gender	Obstruent	Sonorant	Schwa	Full Vowel
Feminine	Subregular: *-(e)n, -s*	Subregular: *-(e)n, -s*	Regular: *-n*	Subregular: *-s, -(e)n*
	Irregular: *-e*	Irregular: *-e*	Irregular: ø	
Masculine	Subregular: *-e, -(e)n, -s*	Subregular: *-e, -(e)n, -s*, ø	Subregular: ø, *-n*	Subregular: *-s, -e*
	Irregular: *-er*	Irregular: *-er*		Irregular: *-er,* ø
Neuter	Subregular: *-e, -(e)n, -s*	Subregular: *-e, -(e)n, -s*, ø	Regular: ø	Regular: *-s*
	Irregular: *-er*	Irregular: *-er*	Irregular: *-n*	Irregular: -er, ø

5. In order to achieve sufficient numbers in each cell, the following simplifications have been made: base-final (fricative and affricate) sibilants have been put together with the other final obstruents, although *–s* suffixation is excluded after sibilants. Word-final central [ə] (= written *–e*) and lower [ɐ] (= written *–er*) of spoken Austrian German have been put together as schwa, and diphthongs have been united with vowels, in both cases despite minor differences in following plural suffixes. Among sonorant-final masculines and neuters zero occurs only if the sonorant is preceded by [ə] (when the [ə] is deleted, the sonorant is syllabic).

According to the system of plural suffixation (plus zero) of Table 3, there is no difference in the distribution after final obstruents and sonorants, except for the cases of sibilants and [ə] followed by sonorant (as mentioned in footnote 6). Starting with feminine nouns, we find, among the productive suffixes, competition between *–en* and much less frequent *–s*, as in *Farm-en* = *Farm-s* 'farm-s' (the reverse distribution after full vowels), whereas *–e* suffixation is unproductive, for example, *Braut*, Pl. *Bräut-e* 'bride-s'. After final schwa, only *–n* is productive, zero occurs unproductively after [ɐ], for example, *Vase-n* 'vase-s', *Mutter-n* 'female screw-s' vs. *Mütter* 'mother-s'.

Masculines and neuters differ only after final schwa: zero is the only productive plural type of neuters, as in *Gebirge* 'mountain range(-s)', whereas *–n* is unproductive (*Auge-n* 'eye-s'). With masculines, productive zero competes with productive *–n* (e.g. *Hase-n* 'hare-s'). Examples for the position after obstruents are the productive masculine types *Quiz-e, Prinz-en, Spot-s* 'quiz-es, prince-s, spot-s' and the unproductive *Wäld-er* 'woods'.[6]

2.3 Danish plural formation

The Danish system of nominal pluralization consists of a number of plural allomorphs, namely the suffixes *a*-schwa, *e*-schwa[7], zero, *-s*, *-a* and *-i*.[8] Among adult plural suffixes (Allan, Holmes and Lundskær-Nielsen 1995 :21–38), the learned suffixes *-a, -i* are irrelevant for our corpus and left out here, and plurals in *–s* occur only marginally in our corpus, for example in *Teletubbies* (in addition to the native form *Teletubbier*). Apart from such English loans, this leaves us with the plural suffixes zero and the two overt suffixes *a*-schwa and *e*-schwa, that is, the two neutral vowels in Danish.[9]

---

6. What is special for the system of oral (Eastern, thus also Viennese) Austrian German is that unstressed word-final orthographic -er is always realised as [ɐ] and thus falls into the cell of word-final schwa and not sonorant. Moreover, in contrast to other types of German, *-n* plurals are productive with masculines and neuters ending in *–l*. Finally, where *–s* plurals compete with other plural patterns, they are less frequent than in Northern Germany.

7. *e*-schwa is a highly assimilable central mid neutral vowel: [ə] (Basbøll 2005: 52-57) and *a*-schwa is a central retracted neutral vowel (a syllabic pharyngeal glide): [ɐ] (Basbøll 2005: 58).

8. Similar to German, the *a*-schwa plural suffix may combine with *Umlaut*, and *Umlaut* can also be the only plural marker (i.e. "combine with zero"). Although the syllable prosody *stød* plays a key role as a cue to morphological structure in Danish (cf. Basbøll 2005: 432-442), in lexical and grammatical respects parallel to tonal word accents in Swedish and Norwegian, it is disregarded in this chapter where only suffixes, not alternations of the base, are considered.

9. There exists a large discrepancy and mismatch between speech and writing in Danish, and there is scarcely any tradition for morphological analysis departing from sound (as against orthography), with the exception of the pronunciation dictionary by Brink, Lund, Heger and Normann Jørgensen, 1991: 1632-1659 (noun plurals are treated on p. 1641-1645). Our morphological analysis, which departs from phonemes rather than letters, results in a completely different system from that found in the standard descriptions.

Danish has two genders, utrum (common) and neuter. The distribution of plural suffixes according to gender and sonority of the base-final phoneme is illustrated in Table 4[10], where the native overt plural suffixes and zero are categorized according to regularity[11] in each of its eight cells (*e*-schwa does not apply to recent loans and thus does not qualify as subregular).

Table 4. Interaction of gender and sonority in Danish[12]

Sonority / **Gender**	**Obstruent**	**Sonorant**	**Schwa**	**Full vowel**
Neuter	Subregular: *-a-schwa, ø*	Subregular: *-a-schwa, ø*	Regular: *-a-schwa*	Subregular: *-a-schwa, ø*
	Irregular: *-e-schwa*	Irregular: *-e-schwa*	Irregular: ø	
Common	Subregular: *-a-schwa, ø*	Subregular: *-a-schwa, ø*	Regular: *-a-schwa*	Subregular: *-a-schwa, ø*
	Irregular: *-e-schwa*	Irregular: *-e-schwa*	Irregular: ø	

The fully mature plural system displayed in Table 4 shows no differences between the two genders in the distribution of plural suffixes according to regularity. However, it is well known from language history that zero plurals are found (relatively) more often in neuters than non-neuters in native simplex words.. There are numerous unambiguous cues for the gender of the singular form of Danish nouns in the linguistic context. A number of such cues within the noun phrase are gender-specific indefinite and definite articles, definite inflectional suffixes of the noun, indefinite inflectional suffixes

10. The columns for base-final sonority of Table 4 make a distinction between glides and vowels, in agreement with the principles of Danish phonology: diphthongs are in all phonological respects VC-sequences (Basbøll 2005: 65-69), as against diphthongs in German, for example. Therefore sonorant consonants and glides are here taken together as constituting the natural sonority class of Sonorant Non-Vowels (cf. Basbøll 2005: 173-201). In relation to choice of plural suffix, *e*-schwa and *a*-schwa are so similar that they are here considered one sonority class of neutral vowels, called Schwa.

11. In addition to productivity, the distribution of plural suffixes in the lexicon has been included in our considerations, but not data from child language acquisition.

12. We gratefully acknowledge the valuable participation of Claus Lambertsen and Laila Kjærbæk Hansen in the work with the tables and on the computational tools used (the OLAM-system), and thank the latter for giving us access to her term paper *Dansk Nominalmorfologi - en empirisk undersøgelse af distributionen af pluralissufikser klassificeret ud fra et lydligt perspektiv i Child directed speech og skreven tekst* (University of Southern Denmark, 2006).

(including zero) of the adjective, and certain pronouns. The question is whether the child can combine information on gender (from singular forms only) with the distribution of plural suffixes, in particular zero.

Radical, partly optional, processes of sound reduction in Danish (Rischel 2003) in many cases obscure the distinction between an overt suffix and zero: for example, plural *bagere* (singular *bager* 'baker') in distinct pronunciation has a (lexicalized) agentive suffix *a*-schwa followed by a plural suffix *a*-schwa, but the difference between one *a*-schwa (in the singular) and two (in the plural) is not at all stable. Thus in reduced speech, there can be complete merger of the singular and plural form, that is, strictly speaking a "zero-plural" rather than the plural suffix *a*-schwa which is found in distinct speech.

A plural suffix in Danish may be followed by an inflectional suffix signalling definiteness and furthermore by a possessive ending (analysed as either a clitic (Herslund 2001, 2002) or an inflectional suffix), for example, *dreng, dreng-e, dreng-e-s, dreng-e-nes* (singular indefinite non-possessive, plural indefinite non-possessive, plural indefinite possessive, plural definite possessive of 'boy'). The fact that the plural suffix in such cases is not word-final would make it more opaque for the language acquiring child than suffixes which always occur at the end of the word (as is the case for overt plural suffixes in the other Germanic languages of this study, definite inflection being a typological characteristic of North Germanic). In the tables on Danish, all noun plurals are analysed together, whether followed by a definite and/or possessive suffix or not.

2.4 Hebrew plural formation

Hebrew is the only Semitic language to participate in this study, and thus its plural system is distinct from the other three languages under investigation here. Hebrew nouns come in two genders – masculine, taking the plural suffix *–im*, and feminine, taking the plural suffix *-ot*. All native Hebrew plurals are formed by suffixation to the final base consonant, with concurrent stress shift to the suffix[13], for example, *tik - tikím* 'bag-s'. Singular masculine nouns are the unmarked form, ending with either a consonant or with the stressed vowel *–e* (e.g., *moré* 'teacher'). Singular feminine nouns end either with the stressed vowel *–a* (e.g., *sirá* 'boat') or with a variety of suffixes all ending with *–t*[14] (*–it* as in *sakít* 'bag'; *-ut* as in *xanút* 'shop'; *-éCet* as in *rakévet* 'train'[15]; *-ot* as in *axót* 'sister'). Nouns ending in a consonant (masculine) attach the plural suffix to the final base consonant (*xatúl - xatul-ím* 'cat-s'). Plural suffixation on nouns ending in stressed *–e* or *–a* replace them with the plural suffix (*more - mor-ím* 'teacher – s', *sirá- sir-ó*t 'boat-s'). Feminine nouns ending in *–t* delete it, attaching plural *–ot* to a *y*-final base (*sakít – sakiy-ót* 'bag / s').

13. Foreign stems do not undergo stress shift.

14. Spelled ת rather than ט.

15. With other allomorphic variations, such as *–áCat* (*caláxat* 'plate').

Table 5. Interaction of gender and sonority in Hebrew

Sonoriy \ Gender	Masculine	Feminine
Obstruent *-t*	Subregular: *–im, –ot*	Regular: *–(y)ot* Irregular: *–im*
Obstruents other than *-t* and sonorants excluding *-n*]	Regular: *–im* Irregular: *–ot*	Subregular: *–im, –ot*
Sonorant *-n*	Subregular: *–im, –ot*	Regular: *–im*
Unstressed *-a*	Subregular: *–im, –ot*	Subregular: *–im, –ot*
Stressed *-e*	Regular: *–im* Irregular: *–ot*	
Stressed *-a*		Regular: *-ot* Irregular: *-im*

The Hebrew-specific manifestation of the sonority scale expresses suffix regularity by the interaction of base-final segments and gender, as shown in Table 5. Masculine stems ending with non-suffixal non-deleting *-t* result in subregular patterns (*sharvit-ím* 'scepter-s', *ot-ót* 'signal-s'); while feminine stems delete suffixal *-t* yielding regular plurals (either replaced by *-y* as in *paxít - paxiyót* 'can-s'; or else, like all other plurals, directly attaching the suffix to the final consonant of the base, as in rakévet – *rakav-ót* 'train-s'). These are followed by masculine stems ending with all other obstruents and sonorants (excluding *-n*), yielding both regular (*pil-ím* 'elephant-s') and irregular suffixes (*kir-ót* 'wall-s'), while such feminine stems yield subregular patterns (*kos-ót* 'glass-es', *cipor-ím* 'bird-s). Masculine stems ending in *–n* (typically *–an* and *–on*) result in subregular patterns (*xalon-ót* 'window-s', *balon-ím* 'baloon-s'), while such feminine stems (which are very scarce) yield regular *ím* suffixation (*éven-avan-ím* 'rock-s'). Stems of both genders ending with an unstressed *-a* (e.g., masculine *ca'acúa - ca'acu'-ím* 'toy-s', feminine *cfardéa- cfarde-ím* 'frog-s') – the latter always actually ending with an underlying "guttural" or pharyngeal – also yield subregular patterns. Finally, stressed *-e* and *-a* yield both regular (masculine *moré - mor-ím* 'teacher-s', feminine *sirá - sir-ót*) and irregular patterns (masculine *mar'é - mar'-ót* 'sight-s' and feminine *nemalá - nemal-ím* 'ant-s').

3. Databases

The analyses presented here are all based on longitudinal recordings of spontaneous samples of speech input to young children and of corresponding children's output in the four languages under investigation. Below, we provide short descriptions of the four language corpora.

3.1 Dutch

The input data reported in this paper are from the Dutch corpora in the CHILDES (MacWhinney 2000) database (http://www.cnts.ua.ac.be/childes/data/Germanic/Dutch/), more specifically the input data to the children Abel, Daan, Iris, Josse, Laura, Matthijs, Niek, Peter, Sarah and Tom, providing information on speech directed to children from the age of 1;05 – 5;06.[16] The exact details concerning data collection, and transcription can be found in the CHILDES database manuals (http://www.cnts.ua.ac.be/childes/manuals/). The children's output data stem from the CHILDES' Dutch triplets corpora (Gijs, Joost, Katelijne and Arnold, Diederik, Maria) and from the unpublished Jolien corpus (Gillis 1997).

3.2 German

The German corpus consists of 137 recordings of two Austrian children aged 1;3 – 6;0 (Jan) and 1;6 – 3;0 (Katharina), audio-recorded at their homes in spontaneous interaction with their mothers. Recording intervals vary from one week (boy Jan from 1;8 – 2;11) to one month in later periods. The data were transcribed, coded and analyzed according to the CHILDES system.

3.3 Danish

The Danish corpus is a small sample of recordings from two Danish twin families, from the Odense Twin Corpus. The two pairs of twins were recorded in their homes in interaction (eating- or playing situation) with their parents or caretaker and the 28 recordings were recorded with intervals of approximately 1 month, when the children

16. Children's age ranges: Abel: 1;10.30 – 3;04.01; Daan: 1;08.21 – 2;03.30; Iris: 2;01.01 – 3;06.15; Josse: 2;00.07 – 3;04.17; Laura: 1;09.04 – 5;06.12; Matthijs: 1;10.13 – 3;07.02; Niek: 2;07.00 -3;10.17; Peter: 1;05.07 – 2;08.22; Sarah: 1;06.16 – 5;02.13; Tomas: 1;07.05 – 3;01.02.

were from age 1;1 to 2;5. The data were transcribed according to the CHILDES system and coded in the OLAM system.[17]

3.4 Hebrew

The study is based on the Berman Longitudinal corpus, 268 audio-recordings containing naturalistic longitudinal speech samples of four Hebrew-speaking children between the ages 1;4 – 3;3.[18] Data consist of spontaneous interactions between the children and their parents. Recording took place in the children's homes, at intervals of approximately 10 days between sessions. Data were transcribed, coded, and analyzed using CHILDES (MacWhinney 2000).

3.5 General frequencies across the four data-sets

Table 6 below presents the information on our four databases. Data is presented in wordforms and tokens, rather than in lemmas, since for this age group, lemmas are too few to really draw conclusions from, while wordforms indicate both lexical and inflectional growth. Also, wordforms cover singulars versus plurals, which is what we are interested in.

Table 6. General word frequencies in types and tokens across the four data-sets

Languages / Frequencies	Dutch Up to 5;6	Austrian German Up to 2;6	Danish Up to 2;5	Hebrew Up to 3;5
Number of word forms in CDS	49,554	6,382	4,384	8,275
Number of word forms in CS	11,868	2,730	1,129	4,142
Number of word tokens in CDS	1,217,341	134,629	117,617	245,384
Number of word tokens in CS	350,543	26,759	13,473	103,226

Our method will consist of identifying noun plurals and characterizing the distribution of noun plural categories in CDS directed to young children learning the four study languages, comparing these data with a similar analysis of the output of those children. We expect to find similar distributional patterns of restrictions in CDS and CS in all

17. The Olam system (developed by Claus Lambertsen, Berlin, and Hans Basbøll and Thomas O. Madsen, Odense) is partly a semi-automatic coding system, which word by word can supply texts in Danish orthography with phonological-/ phonetic-, morphological and segmental information; partly a system, OLAM-search, which can be used for linguistic search purposes, in particular involving phonology, morphology and their interaction.

18. Children's age ranges: Hagar (girl): 1;7–3;3; Lior (girl): 1;5–3;1; Leor (boy): 1;9–3;0; Smadar (girl): 1;4–2;4.

four languages, mediated by the typological differences between Germanic and Semitic languages, on the one hand, and by language-specific differences, on the other.

4. Plurals in child directed speech and child speech

For each language sample, we now present the following data: (i) the number of noun types and tokens in both input and output; (ii) the number of noun plurals in each of these samples, and (iii) their proportion out of all noun types and tokens. Note that we count *types* as *form types* (word forms) rather than word types (lemmas), as more appropriate for the evaluation of early lexical and grammatical development. Thus, Hebrew *tapuz* 'orange' and *tapuzim* 'oranges' would be counted as two types. Proper nouns (=names) were excluded from corpora. Table 7 presents noun and noun plural frequencies in speech directed to young children in various age ranges, up to age 6, with numbers representing the pooled data over all time points and children investigated in each language. These corpora will enable us to trace the changes in noun plural input to older preschoolers, reflecting fine-tuning patterns in parental input to children (Snow 1995).

Across our four languages, between 20% to 24% of the noun types young children are exposed to are noun plural types, while noun plural tokens constitute only between 10% to 15% of the noun tokens they hear. These cross-linguistic data indicate that young children start the route to learning about noun plurals from a small set of noun types and tokens constituting a scant percentage of the nouns they hear.

Table 7. Raw frequencies and percentages of nouns and noun plurals in CDS

Languages / Frequencies	Dutch Up to 5;6	Austrian German up to 6	Danish Up to 2;5	Hebrew[19] Up to 3;5
Number of noun (form) types	8,812	4,009	1,886	2,136
Number of noun tokens	112,732	26,667	9,490	34,671
Number of noun plural (form) types	2,120	871	460	440
Percentage of plural noun (form) types (out of total noun forms)	**24%**	**22%**	**24%**	**21%**
Number of noun plural tokens	16,549	3,600	1,521	3,369
Percentage of plural noun tokens (out of total noun tokens)	**15%**	**14%**	**15%**	**10%**

19. The numbers for Hebrew plural nouns exclude dual nouns, compound nouns (status constructus) in the plural, and Pluralia Tantum.

Table 8. Raw frequencies and percentages of nouns and noun plurals in CS

Frequencies \ Languages	Dutch Up to 3;1	Austrian German Up to 2;6	Danish Up to 2;5	Hebrew Up to 3;5
Number of noun (form) types	2,459	916	439	1,224
Number of noun tokens	14,226	7,007	2,156	21,141
Number of noun plural (form) types	396	142	84	256
Percentage of plural noun (form) types – (out of total noun forms)	**16%**	**16%**	**19%**	**21%**
Number of noun plural tokens	940	549	366	1,635
Percentage of plural noun tokens (out of total noun tokens)	**7%**	**8%**	**17%**	**8%**

Table 8 tells another interesting story, which echoes what we have just seen in the general CDS table: Young children's production of noun plurals in most cases lags somewhat behind that of the input they are exposed to. Thus, in two of our four languages (Dutch and Austrian German), children's noun plural types constitute about 16% of the total noun types, between 5–8% (one third) less than what they hear. Danish and Hebrew-speaking children produce more noun types (around 20%). While the gap between input and output observed for Dutch and German is maintained for Danish (about 5%), the Hebrew data shows no difference in the relative amount of noun plural types. One reason might be the fact that the Hebrew database comes up to age 3;5. Another might be typological – the rich morphological structures of Hebrew may entail earlier learning of morphological types. Regarding noun plural tokens, again three of our four languages show similar patterns of distribution, with about 7% plural tokens in children's output. Here, the Danish data is exceptional, with more than twice as many noun plural tokens.

4.1 Distribution of plural categories in CDS

Having outlined the kind of plural input children hear in Dutch, Austrian German, Danish, and Hebrew, and the kind of plural output they produce in these four languages, we are now ready to proceed to compare the complexity of the mature system with that of CDS and CS. Thus, we next present the distribution of suffixation categories in the speech input to children in each of the languages of our study, by sonority and gender (if the language has gender difference relevant for plural formation).

4.1.1 *Dutch*

Tables 9 and 10 present the analysis of suffix predictability in Dutch CDS. The table is organized in two parts: the figures for noun types are presented in the top panel and

those for tokens in the bottom one. In each panel the two productive suffixes (*-en*, *-s*) are represented, and the results are displayed as absolute figures and as percentages. The table is further organized as follows: separate calculations were carried out for types and tokens regarding what proportion of the words take *–en* respectively *–s* as plural suffix. Thus, for words ending in an obstruent, there were 604 types with *–en* plural and 15 with *–s* plural, and out of 619 word types ending in an obstruent, 97.6% take *–en* as plural suffix, and only 2.4% take *-s*.

Table 9. Suffix distribution on the basis of word-final phonology: types in Dutch CDS

Sonority / Suffix	Consonant						Vowel					
	Obstruent		Sonorant				Schwa		Full vowel			
			Full Vowel + Sonorant		Schwa + Sonorant				Final Stress		Prefinal Stress	
	N	%	N	%	N	%	N	%	N	%	N	%
-en	604	98	366	89	13	6	25	3	19	79	5	6
-s	15	2	44	11	209	94	730	97	5	21	72	94
Total	619		410		222		755		24		77	

Table 10. Suffix distribution on the basis of word-final phonology: tokens in Dutch CDS

Sonority / Suffix	Consonant						Vowel					
	Obstruent		Sonorant				Schwa		Full vowel			
			Full Vowel + Sonorant		Schwa + Sonorant				Final Stress		Prefinal Stress	
	N	%	N	%	N	%	N	%	N	%	N	%
-en	4,827	99.5	3,005	91	68	6	65	1	274	94	25	4
-s	24	0.5	296	9	1,147	94	5,862	99	18	6	668	96
Total	4,851		3,301		1,215		5,927		292		693	

On the whole, the results show that the predictability of the plural suffix in CDS is very high: The token counts all reach a level of more than 90%, and also the type counts indicate predictability of more than 90% (except for one cell: words ending in a full vowel, with final stress). The most straightforward categories are words ending in an obstruent and words ending in a schwa: only the final segment determines the selection of the suffix. Especially for obstruent-final words this comes as a surprise since according to the analysis of the mature system (see section 2), the words' stress pattern plays a role: obstruent-final words with final stress take *-en* and those with penultimate stress take *–s*. However the generalization from CDS is that obstruent-final words take *–en*. Hence only one subregularity from Table 2 is actually represented in CDS.

Informal observation shows that children overgeneralize the use of *–en*: *kok* ('cook') and *jeep* ('jeep') are often pluralized as *kok-en* and *jeep-en* instead of *kok-s* and *jeep-s*.

The choice of the plural suffix in sonorant-final words is also sensitive to the word's stress pattern in the adult system (according to Van Wijk 2002: 87.1% of word tokens with final stress take *–en*, and only 18.0% of word tokens with prefinal stress take *-en*). However in CDS the generalization is somewhat different: if the sonorant is preceded by a full vowel, *-en* is preferred in a majority of cases (tokens: 91%, types: 89%) and when a schwa precedes the sonorant *–s* is predominantly chosen (tokens: 94.4%, types: 94.1%).

The only category in which stress pattern appears to play a role (as in the mature system), are the words ending in a full vowel: when there is final stress, *-en* is the preferred suffix (tokens: 93.8%, types: 79.2%) and words with prefinal stress prefer *–s* (tokens: 96.4%, types: 93.5%).

4.1.2 *German*

The following tables present the analysis of suffix predictability in German CDS in terms of types (Table 11) and tokens (Table 12).[20]

Table 11. Suffix distribution on the basis of item gender and word-final phonology: types in German CDS

Suffix	Sonority / Gender	Obstruent		Sonorant		Schwa		Full Vowel	
		#	%	#	%	#	%	#	%
-s	Feminine			1	4.55			3	50.00
	Masculine	3	4.92	3	5.08			8	61.54
	Neuter							11	57.89
-(e)n	Feminine	5	33.33	20	90.91	100	99.01	2	33.33
	Masculine	7	11.48	7	11.86	3	8.57		
	Neuter	4	7.84	17	26.56	2	22.22		
-e	Feminine	10	66.67	1	4.55			1	16.67
	Masculine	47	77.05	13	22.03			4	30.77
	Neuter	17	33.33	15	23.44				

20. The absolute numbers (both types and tokens) for plurals of nouns ending in obstruents, sonorants and schwa are very similar, which allows to roughly compare percentages in horizontal rows. This is also a reason why we did not introduce word-final sibilants (which block *–s* plural formation) as a separate category: cells for this category would contain rather small numbers but diminish the numbers of word-final obstruent cells, i.e. the numbers of obstruent-final and sonorant-final cells would differ much more.

Suffix	Sonority / Gender	Obstruent		Sonorant		Schwa		Full Vowel	
-er	Feminine								
	Masculine	1	1.64	2	3.39				
	Neuter	29	56.86	4	6.25			4	21.05
zero	Feminine					1	0.99		
	Masculine	3	4.92	34	57.63	32	91.43	1	7.69
	Neuter	1	1.96	28	43.75	7	77.78	4	21.05
N		127		145		145		38	

Table 12. Suffix distribution on the basis of item gender and word-final phonology: tokens in German CDS

Suffix	Sonority / Gender	Obstruent		Sonorant		Schwa		Full Vowel	
		#	%	#	%	#	%	#	%
-s	Feminine			1	1.52			4	36.36
	Masculine	4	1.67	9	3.21			83	82.18
	Neuter							130	86.67
-(e)n	Feminine	5	10.20	64	96.97	331	99.70	5	45.45
	Masculine	32	13.39	20	7.14	15	11.63		
	Neuter	6	2.16	58	23.87	19	44.19		
-e	Feminine	44	89.80	1	1.52			2	18.18
	Masculine	194	81.17	87	31.07			17	16.83
	Neuter	43	15.47	121	49.79				
-er	Feminine								
	Masculine	1	0.42	11	3.93				
	Neuter	225	80.94	5	2.06			15	10.00
zero	Feminine					1	0.30		
	Masculine	8	3.35	153	54.64	114	88.37	1	0.99
	Neuter	4	1.44	59	24.28	24	55.81	5	3.33
N		566		589		504		262	

As in Dutch above, these percentages (calculated in the same way as in the Dutch Tables 9, 10) show clear divergences from what can be found in the German grammars and in the literature on ADS (see Köpcke 1993; Wegener 1999 with references): the plural

suffix -s does not represent the default (as used in the respective claims by Clahsen (1999) and others cited there), because the *–s* plural is highly predictive only for masculine and neuter nouns that end in a full vowel, and its distribution depends clearly on word-final phonology (hardly discussed in the literature, except for vowels (Köpcke 1993 :128–33) and sibilants). Plurals with *–en* are much more of a default for feminines than often assumed in the literature (e.g., Clahsen 1999), not only in the sense of distributional asymmetry, but also in the sense of overall productivity, and there are clear frequency differences between masculines and neuters. The same holds for *–e* plurals, for which also dependency on word-final consonants (obstruents vs. sonorants) is a novel finding. Gender dependency of the distribution of unproductive *–er* plural formation is impressive in its novelty (more than what appears in Köpcke 1993 :109–10, 39–43) as well as the relevance of word-final consonant phonology in neuters.

The various differences between masculine and neuter gender are unexpected, because neuter and masculine inflection are generally considered to belong to the same inflectional classes (Wegener 1999). And in language usage, CDS clearly differs from the mature system in allowing much more predictability. This may also explain why children appear to acquire neuter and masculine gender inflection (Mills 1986) with no greater difficulty than feminine gender (except over-extension of the most frequent definite article form *die* 'Nom. & Acc. Sg. fem. or Pl.'): they are confronted with much less ambiguous signals in CDS than what has been assumed so far.

The only sizable difference between types and tokens is in the much higher token frequency of *–s* plurals after full vowels of neuters. This is due to the frequent use of words like neuter *Auto-s* 'car-s' in CDS to the car-loving boy Jan and to the frequent neuter diminutives in *-i*, Pl. *-i-s,* such as *Has-i-s*, diminutive of *Hase* 'hare', a diminutive type that is restricted to CDS and early CS.

While only 3 of 12 cells in the fully mature system contain regular suffixation, that is a clear default suffix, the table of CDS types contains a greater degree of predictability: 6 cells indicate more than 66,6% predictability of the occurrence of a suffix in a given combination of gender and phonological context, and the occurrence or non-occurrence of a given suffix or zero is highly predictable in at least 40 of 60 cells. The distribution of plural suffixation in CDS can thus be considered to represent the core of plural inflection.

If we compare the distributions in Tables 11 and 12 with the later input (of Jan up to 6;0, of Katharina up to 3;0) then we find little differences: They consist mainly in the filling of some empty slots of the earlier input, but always with very small numbers, so that predictability of non-occurrence decreases only very slightly. Furthermore, sometimes differences in percentages between competing suffixes (in terms of frequency) also decrease, which diminishes the predictability of the dominant competitor.

4.1.3 *Danish*

Table 13. Suffix distribution on the basis of item gender and word-final phonology: types in Danish CDS

Suffix	Gender \ Sonorit	Obstruent		Sonorant		Schwa		Full vowel	
		#	%	#	%	#	%	#	%
a-schwa	Neutrum	4	25	9	20	13	76	7	44
	Utrum	48	62	72	64	127	98	11	52
e-schwa	Neutrum	3	19	2	4	2	12	0	0
	Utrum	18	23	31	28	0	0	0	0
Zero	Neutrum	9	56	35	76	2	12	9	56
	Utrum	11	14	9	8	2	2	10	48
N		93		158		146		37	

Table 13 shows that in five out of the eight gender x sonority combinations there is relatively high predictability, more than 60 %, for the occurrence of one native plural suffix[21] (either *a*-schwa or zero) – a finding which does not follow from the fully mature system displayed in Table 4. In addition, one marker (zero) in the sixth gender x sonority combination (neuters ending in an obstruent) is clearly dominant. For stems ending in a full vowel, *a*-schwa and zero are equally distributed. Only *e*-schwa (which is irregular in the system, see Table 4) is, expectedly, not dominant in any cell. For stems ending in a full vowel or schwa the degree of predictability agrees with the system. But for stems ending in an obstruent, and even more so for stems ending in a sonorant non-vowel, the predictability is clearly higher in CDS than in the system: for neuter nouns zero plurals are dominant whereas for utrum nouns *a*-schwa is dominant. This asymmetrical distribution of *a*-schwa and zero, which adds to the predictability of one suffix in a particular cell, is more clearly seen in the table of tokens, also for bases ending in a full vowel (Table 15).

21. In our CDS corpus the plural suffix *–s* is marginally represented: In addition to the lexical exception *høns* 'hens' (cf. *høner*, '(female) hens', not in our corpus), we have *flutes* (from French *flûtes*, in Danish sometimes, like here in our corpus, pronounced with [s], unlike in French) and *Teletubbies* together with the parallel form *Teletubbier* (plural definite *Teletubbiesene* together with *Teletubbierne,* both in our corpus, cf. 2.3). Opaque plural definite forms like *indianerne* 'the Indians' (cf. 2.3) are represented, but only rarely.

Table 14. Suffix distribution on the basis of item gender and word-final phonology: tokens in Danish CDS

Suffix	Senority / Gender	Obstruent		Sonorant		Schwa		Full vowel	
		#	%	#	%	#	%	#	%
a-schwa	Neutrum	18	25	18	11	51	61	8	13
	Utrum	134	60	208	56	370	99	65	57
e-schwa	Neutrum	3	4	4	2	20	24	0	0
	Utrum	44	20	110	29	0	0	0	0
Zero	Neutrum	50	71	140	86	13	15	53	87
	Utrum	45	20	56	15	3	1	50	43
N		294		536		457		176	

4.1.4 *Hebrew*

Table 15. Suffix distribution on the basis of item gender and word-final phonology: types in Hebrew CDS

Suffix	Sonority / Gender	Obstruent *–t*		Obstruents other than *-t* and sonorants [excluding *-n*]		Sonorant *–n*		Unstressed *-a*		Stressed *-e*		Stressed *-a*	
		#	%	#	%	#	%	#	%	#	%	#	%
-im	Masculine	4	80	208	92	13	52	12	80	4	80	–	
	Feminine	0	0	5	71	1	100	1	17	–		6	5
-ot	Masculine	1	20	18	8	12	48	3	20	1	20	–	
	Feminine	35	100	2	29	–		5	83	–		105	95
N		40		233		26		21		5		111	

In terms of gender, our CDS-sample has 271 masculine, but only 160 feminine noun types – reflecting the historical primacy of masculine *–im* suffixation in Hebrew (Schwarzwald 1983). The largest group of plural types contains nouns on the lower end of the sonority scale, ending with obstruents other than *-t* and sonorants other than *-n* (233 types in total), to which the suffix is directly attached. Table 15 reveals that, within

this category, the most frequent noun plurals are masculine nouns, but also that the most frequent type of suffixation is through the application of the *–im* suffix (for both masculine and feminine nouns). In other words, the bulk of noun plurals with an obstruent or a sonorant excluding *-n* in CDS are inflections of nouns ending with an obstruent, and under both gender conditions, it is highly predictable that such nouns will receive the suffix that is associated with masculine gender – whether such suffixation is regular or subregular in the system. That is, predictability of suffixation is a function of *base-final phonology*. Note, however, that predictability is lower for feminine nouns, in line with their subregular status. In general, these results may explain children's tendency to overgeneralize using the *–im* suffix (Berman 1981; Levy 1980, 1988).

The picture is quite different for the second largest group of noun plurals, 142 noun types ending with the most sonorous vowels as well as the sonorant /n/: For nouns ending with stressed vowels (either *-e* or *-a*), nouns marked for feminine gender consistently take the *–ot* suffix, and nouns marked for masculine gender take *–im* suffixation. And in the case of nouns ending with the sonorant *n* (typically considered a marker of masculine gender), *-im* suffixation is somewhat more predictable, even though their status in the system is subregular. That is, when base-final phonology clearly marks gender, predictability of suffixation not only coincides but is also affected by system regularity.

The third and smallest group (61 types) is nouns ending with *–t* and unstressed *-a*, that is, nouns ending in the obstruent /t/ and in the least sonorous vowel on our scale. Here, it seems that suffixation is crucially dependent on *inherent gender*. Thus, *-im* suffixation is most predictable for masculine nouns ending with the obstruent t, while the *-ot* suffixation is most predictable for their counterpart feminine nouns: indeed, the overwhelming majority of nouns ending with the obstruent *t* are feminine, as clearly shown by the higher number of types in this cell (35). Similarly, predictability of suffixation for nouns ending with an unstressed *-a* vowel is also determined by gender – with an 80% chance of *–im* plurals being masculine nouns and 83% chance of *–ot* plurals being feminine nouns.

These results are not only strikingly similar but even more pronounced when considering noun plural tokens:

Thus, for example, 81% of all feminine noun tokens ending with obstruents other than *-t* and sonorants other than *n* receive the *–im* suffix (as compared to 71% of the same nouns in terms of types); 96% of all nouns ending with stressed *-e* take *–im* suffixation (as compared to 80% in terms of types); predictability of *-im* suffixation for nouns ending with the sonorant *-n* is much higher (79% of all tokens as compared to 52% of all types); and for nouns ending with an unstressed *a* vowel, there's a 95% (as compared to 80%) chance of *–im* plurals being masculine nouns.

Table 16. Suffix distribution on the basis of item gender and word-final phonology: tokens in Hebrew CDS

Suffix	Gender \ Sonority	Obstruent *-t*		Obstruents other than *-t* and sonorants [excluding *-n*]		Sonorant *-n*		Unstressed *-a*		Stressed *-e*		Stressed *-a*	
		#	%	#	%	#	%	#	%	#	%	#	%
-im	Masculine	10	91	1604	92	113	79	255	95	21	96	–	
	Feminine	–		65	81	6	100	4	14	–		9	1
-ot	Masculine	1	9	150	8	30	21	13	5	1	4	–	
	Feminine	180	100	15	19	–		25	86	–		775	99
N	3278	191		1834		150		297		22		784	

Our application of the novel gender x sonority interaction to Hebrew plural suffixation has yielded two interesting insights. Firstly, it enabled us to uncover the core of the noun plural system as it is presented to children in CDS, which looks very different from the mature system: Masculine nouns have a much larger representation than do feminine nouns, and most nouns, whether masculine or feminine, take regular suffixation. Subregularities are almost absent from CDS plurals. These characteristics of the core plural system of Hebrew have never been outlined before. Secondly, our analysis also reveals that the distribution in the core system directs Hebrew-speaking children to adhere to two cues – suffixation following base-final phonology on the one hand, and suffixation following inherent gender on the other. These cues will enable them later on to untangle subregularities when the core system is extended to its more complex, mature version.

4.2 Distribution of plural categories in CS

To consider the relationship between noun plurals in the input to and output of young children, we now present the same information as in section 4 above in children's output, by sonority and gender: here we are restricted to three languages – German, Danish, and Hebrew.

4.2.1 *German*

Clear similarities and differences emerge in the comparison of the output (Tables 17 and 18) to the respective input tables: the *–s* plural tokens are much higher in the

output, again due to Jan's predilection of *Auto-s*, *-en* plurals are more frequent in the output, reflecting their typical role in early overgeneralisation (Klampfer and Korecky-Kröll 2002; Sedlak, Klampfer, Müller and Dressler 1998; Vollmann, Sedlak, Müller and Vassilakou 1997). Zero plurals are less frequent in the output: one possible reason is children's preference for iconic suffixation over non-iconic zero marking (Korecky-Kröll and Dressler in preparation). A second reason might be under-representation of zero plurals in children's output where, due to rigorous exclusion of ambiguous forms, some zero plurals may have been counted as singulars.

Table 17. Suffix distribution on the basis of item gender and word-final phonology: types in German CS

Suffix	Sonority / Gender	Obstruent		Sonorant		Schwa		Full Vowel	
		#	%	#	%	#	%	#	%
-s	Feminine								
	Masculine							1	25.00
	Neuter					1	25.00	7	63.64
-(e)n	Feminine			7	87.50	31	100.00		
	Masculine	6	35.29	4	19.05	2	25.00	1	25.00
	Neuter	1	6.67	5	31.25	1	25.00		
-e	Feminine	6	85.71	1	12.50				
	Masculine	10	58.82	7	33.33			1	25.00
	Neuter	6	40.00	7	43.75			1	9.09
-er	Feminine								
	Masculine			4	19.05				
	Neuter	8	53.33					2	18.18
zero	Feminine								
	Masculine	1	5.88	5	23.81	6	75.00	1	25.00
	Neuter			4	25.00	2	50.00	1	9.09
*-en+U	Feminine	1	14.29						
	Masculine			1	4.76				
	Neuter								
N		39		45		43		15	

Table 18. Suffix distribution on the basis of item gender and word-final phonology: tokens in German CS

Suffix	Sonority / Gender	Obstruent		Sonorant		Schwa		Full Vowel	
		#	%	#	%	#	%	#	%
-s	Feminine								
	Masculine							47	81.03
	Neuter					1	3.85	84	84.85
-(e)n	Feminine			15	93.75	91	100		
	Masculine	6	18.18	6	10.34	3	16.67	1	1.72
	Neuter	1	1.69	30	47.62	22	84.62		
-e	Feminine	25	89.29	1	6.25				
	Masculine	26	78.79	28	48.28			8	13.79
	Neuter	10	16.95	24	38.10			1	1.01
-er	Feminine								
	Masculine			8	13.79				
	Neuter	48	81.36					13	13.13
zero	Feminine								
	Masculine	1	3.03	14	24.14	15	83.33	2	3.45
	Neuter			9	14.29	3	11.54	1	1.01
*-en+U	Feminine	3	10.71						
	Masculine			2	3.45				
	Neuter								
N		120		137		135		157	

Moreover, there are more empty cells in the output than in the input, which we interpret as children ignoring infrequent plural types in the input. Cases in point are *–s* plurals except after word-final full vowels and *-en* plurals after feminine nouns ending in obstruents. The greatest differences are in the distributions after word-final consonants: the children produce illegal umlauted *–en* plurals instead of feminine unproductive umlauted *–e* plurals (which are productive with masculines and neuters) or productive non-umlauted *–en* plurals. Thus they do not seem to grasp, at first, the mutual relevance of word-final phonology and gender in these distributions. After 2;6, Jan and Katharina cease to produce illegal umlauted *-en* plurals, whereas they continue to produce potential but non-existing umlauted *-e* plurals. We interpret this change as indicating that by then they have grasped an important property of core morphology.

4.2.2 *Danish*

Table 19. Suffix distribution on the basis of item gender and word-final phonology: types in Danish CS

Suffix	Gender \ Sonority	Obstruent		Sonorant		Schwa		Full vowel	
		#	%	#	%	#	%	#	%
a-schwa	Neutrum	1	33	1	17	2	67	1	33
	Utrum	7	50	11	53	24	100	4	57
e-schwa	Neutrum	0	0	0	0	1	33	0	0
	Utrum	5	36	10	48	0	0	0	0
Zero	Neutrum	2	67	5	83	0	0	2	67
	Utrum	2	14	0	0	0	0	3	43
N		17		27		27		10	

Comparing the output and the input tables we see a similar pattern in general with a distributional asymmetry after consonants between *a*-schwa and zero, depending on gender. Moreover, for nouns ending in a full vowel, zero plurals are strongly represented even in utrum nouns, in particular in token frequency (more so than in CDS).

Table 20. Suffix distribution on the basis of item gender and word-final phonology: tokens in Danish CS

Suffix	Gender \ Sonority	Obstruent		Sonorant		Schwa		Full vowel	
		#	%	#	%	#	%	#	%
a-schwa	Neutrum	1	25	2	12	14	93	4	29
	Utrum	68	78	20	54	120	100	11	28
e-schwa	Neutrum	0	0	0	0	1	7	0	0
	Utrum	15	17	17	46	0	0	0	0
Zero	Neutrum	3	75	15	88	0	0	10	71
	Utrum	4	5	0	0	0	0	29	73
N		91		54		135		54	

To illustrate the pattern of productivity of the endings, we found, in a particular subcorpus[22], only one instance of an overgeneralization of the plural suffix *e*-schwa (**fiske* for the zero plural *fisk* 'fish'). In all other cases either *a*-schwa or zero were overgeneralized (e.g. **abekatter* for *abekatte* 'monkies' and **gulerød* for *gulerødder* 'carrots', respectively).

4.2.3 *Hebrew*

Table 21. Suffix distribution on the basis of item gender and word-final phonology: types in Hebrew CS

Gender	Sonority	Obstruent *-t*		Obstruents other than *-t* and sonorants [excluding *-n*]		Sonorant *-n*		Unstressed *-a*		Stressed *–e*		Stressed *-a*	
		#	%	#	%	#	%	#	%	#	%	#	%
-im	Masculine	2	100	117	92	12	71	3	50	1	50	–	
	Feminine	1	6	5	83	1	100	1	25	–		7	11
-ot	Masculine	–		10	8	5	29	3	50	1	50	–	
	Feminine	16	95	1	17	–		3	75	–		60	89
N	249	19		133		18		10		2		67	

Hebrew child speech closely reproduces the system as it is presented to children in CDS. All of the phenomena described above characterize plurals produced by children: most noun plurals are masculine and take the suffix *–im*, followed by a much smaller group of feminine nouns marked by *–a* and *–t*, taking the regular feminine suffix *–ot.* Children are thus shown to faithfully adhere to the strongly predictable and regular characteristics of the Hebrew core plural system.

22. The subcorpus consists of one of the twin pairs in our main corpus, ages 2;6-5;8, only common nouns (1226 tokens) and proper nouns (233 tokens) are transcribed and analysed.

Table 22. Suffix distribution on the basis of item gender and word-final phonology: tokens in Hebrew CS

Gender \ Sonority		Obstruent *–t*		Obstruents other than *-t* and sonorants [excluding *-n*]		Sonorant *-n*		Unstressed *-a*		Stressed *–e*		Stressed *-a*	
		#	%	#	%	#	%	#	%	#	%	#	%
-im	Masculine	5	100	783	92	79	87	77	92	1	50	–	
	Feminine	1	1	48	98	7	100	2	23	–		16	4
-ot	Masculine	–		72	8	12	13	7	8	1	50	–	
	Feminine	69	99	1	2	–		7	77	–		349	96
N	1267	75		904		98		93		2		95	

5. General discussion

Our study has focused on noun plural formation, a central area of inflectional morphology, as transmitted by care-takers to young children from birth to the middle of their third year of life. For each of the four languages we investigate – Dutch, Austrian German, Danish, and Hebrew – we have shown two important and novel findings. First, we have shown that quantitatively, children's plural output is closely paced by the input they receive. The amount of noun plurals in speech addressed to children is rather low – about 20% of all noun types and 10% of noun tokens are plural (increasing to about 23% and 14% respectively in CDS of the two Austrian children of this study); and this ratio is closely echoed by the ratio of noun plurals in the output of those very children exposed to the speech we analyzed: about 16% plural types and 7% plural tokens, rising to 17,5% (types) and 11,8% (tokens) of the Austrian children in the period 2;7 – 3;0. This is the first time such a close quantitative relationship has been shown to exist between input and output of plurals.

A second major finding of this paper is qualitative, and provides a first window on what we term *core morphology*. Section 1 discussed the complex interface of gender and sonority in determining suffix predictability, while in Section 2 we demonstrated specifically how this interface generates the complex plural systems of the three Germanic languages and the Semitic language under consideration. Examining the distribution of noun plurals in the longitudinal data of children and their caregivers, our second novel finding is to what extent the complex full adult plural systems described in section 2 above differs from the systems presented to children in the distribution of nouns in the cells created by the intersection of sonority and gender. In all four

languages, our analyses reveal surprising distributions when compared to the fully mature systems adult system. We have found, for all four languages, that plural suffixes directed to children are much more predictable and regular than in the fully adult mature systems, while regularities are given salient, prominent proportions and therefore support children's first forays into the system.

The Dutch analysis thus shows that plural suffixes in CDS are very highly predictable, and that final segments determine suffix selection much more than does the stress pattern. Only one subregularity (out of three) is represented as default/clearly dominant for Dutch in each phonological environment. In the same way, the German analysis resulted in novel findings regarding each of the plural suffixes, showing that *–en* or *–e* plurals rather than *–s* plurals are the default whenever there is a clear dominance of one suffix, links with word-final phonology in *–e* plurals, and interesting interactions with gender. Again, as in Dutch, suffix predictability pervades the child-directed system. In Danish, zero plurals and *a*-schwa plurals after consonants seem to have a more complementary distribution, dependent on gender, and thus a higher predictability, in CDS than in the adult mature system. The complex Hebrew plural system is reduced in CDS mostly to masculine nouns predictably taking the masculine *–im* plural suffix, with regular suffixation of both masculine and feminine nouns. All of these qualitative patterns are echoed in children's output as analyzed in our work.

5.1 CDS compared with adult directed speech (ADS)

While the difference between the plural systems described in Section 2 and CDS is eminently clear, it does not represent a difference between the speech directed to children versus the speech directed to adults. In order to gain an insight into the regularities of plural formation in adult directed speech (as opposed to child directed speech), and more specifically in order to compute the predictability of the plural suffixes in ADS, we needed to consult a database of spoken adult usage. Of the four languages under investigation, only Dutch has such an appropriate corpus. The Spoken Dutch Corpus[23] was consulted. This corpus of approximately 10 million words of contemporary spoken Dutch, collected around the turn of the 21st century, consists of a variety of discourse types (spontaneous conversations, face-to-face as well as over the telephone, lectures, radio and television broadcasts, etc.), which is stratified socially as well as geographically. Due to legal restrictions, the participants were all at least 18 years of age. Hence, this corpus is a genuine sample of adult-directed spoken language.

The corpus is completely part-of-speech tagged and thus represents a highly rich source of data. 998,046 tokens of nominals were identified (excluding proper nouns), representing a rough 10% of the entire corpus, of which 213,699 (21.4%) nominal plural tokens (23,319 plural types). The distribution of the suffixes is as follows: 59.6% of all types take *–en*, 38.8% take *–s*, 0.4% take *–eren* and 1.2% take another suffix. And

23. http://www.tst.inl.nl/cgndocs/doc_English/start.htm

for tokens: 71.6% *-en*, 25.3% *-s*, 2.3% *-eren*, and 0.7% another suffix. The latter two categories will not be considered in what follows.

When we compute the distribution of the plural suffixes according to the phonological form of the singular, similar to Tables 9 and 10 for CDS, it appears that plural formation is highly predictable in ADS. Figures I (types) and II (tokens) compare the predictability of the plural suffix –en in Dutch ADS and CDS according to the form of the final rhyme.

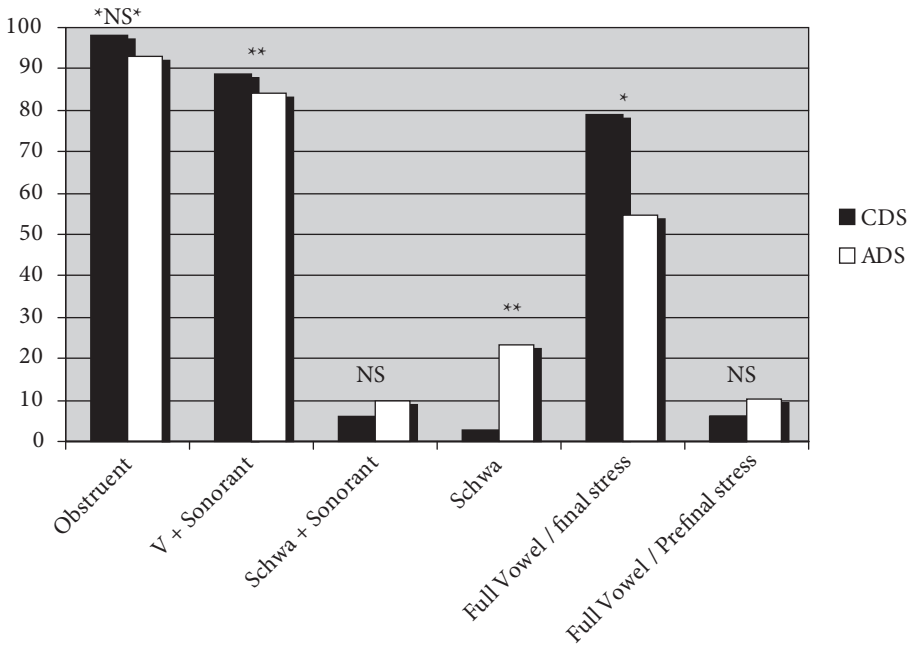

Figure I. Predictability of the plural suffix –en in Dutch ADS and CDS according to the form of the final rhyme (wordtypes)

Figure I clearly shows that in ADS the suffix *–en* (and consequently also the suffix *–s*) is indeed highly predictable, yet is slightly less predictable than in CDS. For instance, wordtypes ending in an obstruent take *–en* as a suffix in 97.6% of the cases in CDS, while in only 93.0% of the cases in ADS (the levels of statistical significance are indicated in Figure I: ** = $p<0.01$, * = $p<0.05$, NS = $p>0.05$). It appears that except for two categories of words for which the difference is only marginally significant, plural formation in ADS is significantly less predictable than in CDS. In other words, while predicatbility of the suffix is high in adult speech, it is even higher in CDS.

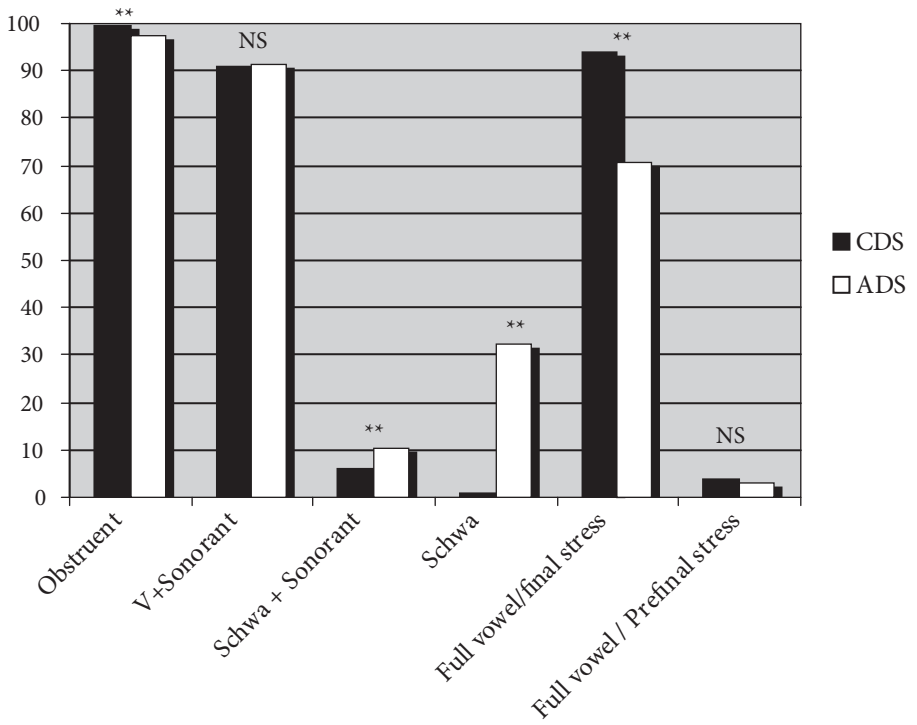

Figure II. Predictability of the plural suffix –en in Dutch ADS and CDS according to the form of the final rhyme (wordtokens)

In terms of wordtokens, Figure II shows, again, that the predictability of the plural suffix is particularly high for all kinds of words. And yet again, CDS is even more predictable than ADS, except for two classes of words, viz. words ending in a full vowel plus a sonorant, and words ending in a full vowel that have prefinal stress. But the difference in predictability between CDS and ADS is not statistically significant for those categories.

Note that the major differences in predictability are to be found in words ending in a schwa and words ending in a full vowel with final stress. The latter category is not surprising: Dutch words ending in a full vowel are typically loans from Romance origin, which are not part of the CDS register. Words ending in a schwa are typically part of CDS, however, while a great majority of these words are diminutivized nouns in CDS (Gillis 1997), the proportion of diminutives is much lower in ADS: in our corpus of ADS only 5% of all nouns are diminutivized, and only 1% of all nouns are pluralized diminutives.

We have thus shown that under the same circumstances of production (speech), CDS has enhanced predictability compared with ADS. While this has been shown so

far only for Dutch, we expect future analyses of spoken corpora in other languages to reveal the same results.

5.2 Typological perspectives

As expected from our previous work (Stephany 2002; Laaha and Gillis 2007), morphological language typology has an impact on the acquisition of core morphology via input to young children. Thus greater morphological richness has been found to stimulate children to acquire inflectional morphology more rapidly than a poorer morphological input system. As Gillis and Ravid (2006) demonstrate, children growing up in a language with a rich morphology carry over such morphologically based strategies even to written language.

If neither gender nor word-final phonology conditions the choice of plural suffixation, as is the case in Turkish, or when word-final phonology predicts plural allomorphy in a purely phonological way, as in English, we do not expect any morphological difference between CDS and ADS. When word-final phonology but not gender conditions the selection of plural suffixes in a phonologically arbitrary way, as in Dutch, then core morphology has been found to be more predictive than the adult system, due to more and stronger asymmetries in the distribution of plural suffixes. When, in addition to word-final phonology, overt gender differences are relevant for the selection of plural suffixes, then CDS also contrasts genders in a more predictive way, as in Danish and in Hebrew with its richer morphology. When even three genders are distinguished, as in German, then CDS even differentiates masculine and neuter gender in its impact on plural suffixation beyond the adult system. We would expect similar phenomena in Slavic languages, where the inflectional morphology of neuters and masculines is very similar as well. In Laaha and Gillis (2007) we established that the richer adult morphology, the speedier children tend to acquire it. A related effect has been found in this study, namely that Hebrew, the richest morphology of our languages appears to stimulate children to produce the highest percentage of plural types.

6. Conclusions

Input plurals, as identified and analyzed in this work, have been found to be simpler, more predictable and thus easier to acquire than the adult systems of plural formation as described in grammars. Plural formation in CDS is generally simpler than in the adult system in avoiding learned plurals and alternative plural variants of the same lexical entry or with the same base phonology. Third and most important, the dependence of the distribution of plural suffixation on gender and on the phonology of the right edge of lexical bases is much more predictable in CDS than follows from adult grammar.

Where do these differences come from? What is the source of the discrepancy between the full adult systems characterized with much irregularity and unpredictability, on the one hand, and the simpler, more regular and more predictable plurals addressed to children, on the other? More data and more analyses are needed to answer this question following the novel findings revealed in this cross-linguistic study. However, we can already point at some directions. It makes sense that singular and plural nouns occurring in the speech directed to children mostly refer to those concrete objects in the child's vicinity which are perceptually salient. Finally, the plurals used in CDS might reveal strong statistical tendencies inherent in each of the languages under investigation, in a sense, the core of each system, which is expanded and elaborated in later language development. Thus in the future, it remains to be investigated to what extent the pragmatic and semantic character of plural nouns addressed to children is related to their formal inflectional features.