



**Asymmetries in the Acquisition of Subject-Verb Agreement  
in Dutch?  
Evidence from Comprehension and Production**

Journal:	<i>First Language</i>
Manuscript ID:	FLA-14-0003.R2
Manuscript Type:	Original Manuscript
Keywords:	subject-verb agreement, number, Dutch, production, comprehension, picture selection, sentence completion, phonological salience, cue reliability
Abstract:	<p>Across languages, children do not comprehend 3SG/3PL subject-verb agreement before age five, despite early mastery in spontaneous speech. This study investigates subject-verb agreement in a language hitherto not studied in this respect, namely Dutch. We examine if (1) Dutch two- and three-year-olds comprehend subject-verb agreement and (2) a comprehension-production asymmetry still exists if task materials are kept constant across domains. Dutch-speaking two- and three-year-olds completed a comprehension (picture selection) and production (sentence completion) task testing 3SG and 3PL. In comprehension, both groups performed above chance on 3PL but not on 3SG. In production, accuracy on 3PL and 3SG was slightly higher than in comprehension. Comprehension and production were moderately correlated. These results show that comprehension is earlier in Dutch than in previously investigated languages, but only for 3PL, suggesting that phonological salience and cue reliability are important. They also show an asymmetry between comprehension and production, albeit much smaller than assumed in previous studies.</p>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

SCHOLARONE™  
Manuscripts

For Peer Review

## Asymmetries in the Acquisition of Subject-Verb Agreement in Dutch?

## Evidence from Comprehension and Production

Across languages, it takes until the age of five before children use subject-verb agreement as a cue to resolve singular-plural contrasts in comprehension. Specifically, studies on English-, Spanish-, Xhosa- and German-speaking children show that children younger than five do not perform above chance when matching third person singular (3SG)- and third person plural (3PL)-sentences to pictures of one versus multiple referents (Johnson, de Villiers, & Seymour, 2005 for English; Pérez-Leroux, 2005 for Spanish; Gxilishe, Smouse, Xhalisa, & de Villiers, 2009 for Xhosa; Brandt-Kobebe & Höhle, 2010 for German). These findings stand in sharp contrast with observations from spontaneous speech that three-year-old children reach high accuracy on subject-verb agreement marking in production (Brown, 1973 for English; Montrul, 2004 for Spanish; de Villiers & Gxilishe, 2008 for Xhosa; Clahsen, 1986 for German; Leonard et al., 1992 for Italian). In this study, we investigate the comprehension of subject-verb agreement in a language hitherto not studied in this respect, namely Dutch. We target two questions. First, we ask if Dutch-speaking two- and three-year old children are able to comprehend subject-verb agreement. Second, we ask if there is evidence for an asymmetry between comprehension and production in the acquisition of subject-verb agreement in Dutch.

**Cross-Linguistic Differences in the Comprehension of Subject-Verb Agreement**

Previous studies have investigated when children grasp the distinction between singular and plural agreement in various languages. For English, Johnson et al. (2005) conducted a picture-selection experiment, presenting children with sentences in which the only cue for either a singular or plural meaning was the agreement suffix on the verb (e.g., *The duck swims on the pond* vs. *The ducks swim on the pond*). In these sentences, the plural

1  
2  
3 morpheme on the noun is masked due to co-articulation effects between the noun and the  
4  
5 verb, so the only disambiguation cue is the presence or absence of ‘-s’ on the verb indicating  
6  
7 either a 3SG or a 3PL reading. Johnson et al. found that five- and six-year-old children  
8  
9 performed above chance on picture-selection, but three- and four-year-old-children did not.  
10  
11 Accuracy was higher on 3SG than on 3PL in all age groups. To explain children’s poor  
12  
13 overall performance, the authors propose that the English inflectional system is too weak to  
14  
15 provide a cue.  
16  
17

18  
19 Pérez-Leroux (2005) replicated Johnson et al.’s experiment in Spanish. Spanish has a  
20  
21 richer inflectional paradigm compared to English and has pro-drop. Despite these differences  
22  
23 between Spanish and English, Pérez-Leroux found very similar results: Performance was  
24  
25 above chance in children aged between 4;8 and 6;6, but not in children between three- and  
26  
27 4;5-years. This is further evidence that the comprehension of number agreement is late, even  
28  
29 in a language with a rich inflectional paradigm like Spanish in which subjects are often  
30  
31 covert. A difference between English- and Spanish-speaking children was also found:  
32  
33 Whereas English children performed better on 3SG than on 3PL, Spanish children performed  
34  
35 better on 3PL than on 3SG. According to Pérez-Leroux, this indicates that children acquire  
36  
37 overt before null morphemes.  
38  
39

40  
41 Late comprehension was also found in German, a language that expresses both 3SG  
42  
43 and 3PL with overt morphemes. Brandt-Kobele and Höhle (2010) tested three- to four-year-  
44  
45 old German-speaking children on a picture selection task similar to the task used in Johnson  
46  
47 et al. and Pérez-Leroux. Test sentences contained the personal pronoun *sie* that is  
48  
49 homophonous between a 3SG (feminine) and 3PL meaning, thereby leaving verbal number  
50  
51 agreement as the only disambiguation cue (e.g., *Sie füttert einen Hund* ‘She feeds a dog’ vs.  
52  
53 *Sie füttern einen Hund* ‘They feed a dog’). Brandt-Kobele and Höhle found that German  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 three- to four-year olds did not perform above chance level in a picture selection task, on  
4  
5 neither of the two morphemes.  
6

7 This review of studies shows that, cross-linguistically, children younger than five are  
8  
9 unable to use number agreement as a cue in comprehension. This holds for English and  
10  
11 Spanish in which either 3SG or 3PL is expressed with a null morpheme as well as German in  
12  
13 which both morphemes are overt. It was further replicated for Xhosa, a language in which  
14  
15 both morphemes are overt and expressed through prefixation (Gxilishe et al., 2009).  
16

17  
18 An exception to these studies showing late comprehension of number agreement is a  
19  
20 study on French by Legendre, Barrière, Goyet, and Nazzi (2010). These authors tested French  
21  
22 children on their comprehension of the singular-plural contrast in a preferential looking and a  
23  
24 (manual) selection experiment. In the selection experiment, 24- and 30-month-old French  
25  
26 children were presented with sentences containing the phrases *Il embrasse* ‘He kisses’ and *Ils*  
27  
28 *embrassent* ‘They kiss’. The crucial contrast between these phrases involves the  
29  
30 pronunciation of the subject pronouns. In French, the 3SG and 3PL subject pronouns are  
31  
32 normally pronounced in the same way (i.e., /il/), but when the following verb is vowel-initial,  
33  
34 the final consonant of the 3PL pronoun (/z/) is pronounced (a phenomenon called ‘liaison’).  
35  
36 Following others, Legendre et al. assume that /z/ is situated in the verb’s onset and thus  
37  
38 should be analyzed as an agreement marker. Their aim was to investigate if young French  
39  
40 children are sensitive to this singular-plural contrast in comprehension. They found that 30-  
41  
42 month-olds – but not 24-month-olds – performed above chance on both 3SG and 3PL.  
43  
44  
45  
46

47 To explain French children’s precocious comprehension, Legendre, Culbertson,  
48  
49 Zaroukian, Hsin, Barrière and Nazzi (2013) propose an account in terms of  
50  
51 morphophonological differences across languages. Specifically, they argue that two factors  
52  
53 play a role: perceptual salience and cue reliability of agreement morphemes. As for perceptual  
54  
55 salience, Legendre et al. (2013) cite work on speech perception suggesting that /z/ in French  
56  
57  
58  
59  
60

liaison is more salient to detect than Spanish /n/, that is in turn more salient than English /s/. As for cue reliability, or the extent to which there is a transparent, one-to-one relationship between a morpheme and its function, several researchers have proposed this as a factor explaining both within and across language differences in the acquisition of morphemes (cf. Brown, 1973; MacWhinney et al., 1984). Cue reliability can explain why number agreement is later in English and Spanish than in French: English ‘-s’ is opaque, because the same morpheme marks plural on nouns, and is also used for the possessive. Spanish ‘-n’ does not provide a very reliable cue either, as it marks 2PL but not 1PL and often occurs in nouns and verbs. French /z/, in contrast, is very reliable, because it is extremely rare as an initial consonant in French words, and marks plural on both nouns and verbs (Legendre et al., 2013).

### **Comprehension-Production Asymmetry**

Production studies show that, across languages, children acquire subject-verb agreement at a relatively young age in spontaneous speech. For English, Brown (1973) found 90% correct use of 3SG ‘-s’ in obligatory contexts between 26 and 46 months of age (Brown, 1973). In Spanish, number morphemes emerge very early in children’s spontaneous speech, around 1;6-1;7 (Montrul, 2004) and may be productive in children as young as 1;10-2;2 (Gathercole, Sebastián, & Soto, 2002). German-speaking children reach 90% accuracy on 3SG and 3PL in spontaneous speech at three years (Clahsen, 1986), or even earlier, around age two (Poeppel & Wexler, 1993).

This high production accuracy at early stages of acquisition is surprising in light of children’s late comprehension of subject-verb agreement. Yet, it may be premature to conclude that there is an asymmetry between comprehension and production in number agreement acquisition: None of the previous studies have directly compared comprehension and production in the same children. Rather, previous claims are based on comparisons between comprehension data, collected through picture selection, and production data from

1  
2  
3 spontaneous speech. Spontaneous speech is likely to overestimate children's agreement  
4  
5 marking abilities, however, as it typically contains familiar and high-frequent verbs on which  
6  
7 subject-verb agreement emerges first (Rubino & Pine, 1998; Wilson, 2003). Evidence that  
8  
9 spontaneous speech data may present an overestimate comes from Brandt-Kobele and  
10  
11 Höhle's (2010) study on German children. These authors collected parental report data on  
12  
13 children's production of 3SG and 3PL morphemes and found that in their sample of three- to  
14  
15 four-year-olds, 72% of the children showed productive use of 3SG and only 48% showed  
16  
17 productive use of 3PL, questioning the idea of an asymmetrical development between  
18  
19 comprehension and production. In a preferential looking study using the same materials as the  
20  
21 picture selection experiment, these authors also found that German three- and four-year-olds  
22  
23 looked longer to the matching pictures in both a 3SG and 3PL condition. This suggests that  
24  
25 they had some knowledge of subject-verb agreement that, for some reason, did not show up in  
26  
27 their selection of the pictures through pointing. The authors argue that comprehension-  
28  
29 production asymmetries may result from diverging task demands rather than from genuine  
30  
31 differences in children's comprehension or production abilities. Specifically, spontaneous  
32  
33 speech may overestimate children's production abilities as compared to production tasks or  
34  
35 parental report, while comprehension experiments using picture selection may underestimate  
36  
37 children's comprehension abilities as compared to looking experiments in which no manual  
38  
39 (pointing) responses are required.  
40  
41  
42  
43  
44

#### 45 **This Study: Subject-Verb Agreement in Dutch**

46  
47 In this study, we examine the comprehension and production of subject-verb  
48  
49 agreement in Dutch. Dutch is different from English, Spanish and French in that both 3SG  
50  
51 and 3PL are marked with an overt agreement suffix. It is similar to German, except that the  
52  
53 3PL morpheme is more salient in Dutch, because the Dutch plural morpheme adds an  
54  
55 additional syllable to the verb stem whereas the German plural morpheme does not (e.g.,  
56  
57  
58  
59  
60

Dutch *voer-en* vs. German *fütter-n*, ‘feed’). The sentences in (1) and (2) illustrate Dutch 3SG and 3PL agreement. An overview of the whole paradigm is given in Table 1.

- (1) Het meisje drink-t ‘The girl drink-3SG’
- (2) De meisjes drink-en ‘The girls drink-3PL’

[Insert Table 1 about here]

In terms of phonological salience, the 3PL form is more salient, due to its extra-syllabic character, than 3SG ‘-t’. Cue reliability is also stronger for 3PL than for 3SG. First, 3PL is used to mark all plural forms in the paradigm, whereas 3SG is also used for 2SG but not for 1SG. Second, the plural suffix ‘-en’ also marks plurality on nouns (e.g., *fiets-en* ‘bicycles’), but there is no such additional ‘singularity function’ for 3SG. For both morphemes, additional functions are found in Dutch. The ‘-en’ suffix occurs on infinitives (e.g., *surf-en* ‘surf’) as well as strong participles (e.g., *gelopen* ‘walked’), and the ‘-t’ suffix also occurs on regular past participle forms such as (e.g., *gesurf-t* ‘surfed’). In this respect, Dutch differs from French in which /z/ is only used to express a plural meaning, and does not have other functions.

Dutch-speaking children start to use subject-verb agreement more productively around age 2;6 in spontaneous production; that is, from this age they contrast different agreement markings and use agreement marking with a wider variety of verb types (Blom & Wijnen, 2013). Evidence from elicited production shows highly accurate and productive use of 3SG and 3PL at age three (Polišenská, 2010). Specifically, Polišenská (2010) tested twelve Dutch three-year-olds in an elicitation task in which children were asked to finish prompt sentences



1  
2  
3 by the experimenter. She found that accuracy scores were 93% and 100% on 3SG and 3PL  
4 marking of existing verbs and 100% and 93% on 3SG and 3PL marking of nonsense verbs.  
5  
6  
7 This high performance on both existing and nonsense verbs indicates that children's  
8  
9  
10 production of 3SG and 3PL agreement suffixes is productive at three years. However, in this  
11  
12 study, only two existing verbs (*tekenen* 'draw' and *drinken* 'drink') and two nonsense verbs  
13  
14 were included, and a relatively small sample of three-year olds participated ( $N = 12$ ). Previous  
15  
16 studies have shown that children's first productions of subject-verb agreement are often  
17  
18 lexically specific and subject to individual differences (Gathercole et al., 2002; Wilson, 2003).  
19  
20 So, a larger sample of Dutch children and a larger set of different verbs are needed to  
21  
22 investigate Dutch children's production of subject-verb agreement. In addition, in order to  
23  
24 investigate which morpheme is acquired first (3SG or 3PL), younger children need to be  
25  
26 tested than in Polišenská (2010), to avoid ceiling effects.  
27  
28

29  
30 To date, no studies have investigated when Dutch children develop an understanding  
31  
32 of verbal number agreement, and none of the previous cross-linguistic studies on the  
33  
34 comprehension of number agreement have directly compared children's comprehension and  
35  
36 production of verbal number agreement using the same task materials.<sup>1</sup> The aim of this study  
37  
38 is twofold. First, we investigate Dutch children's comprehension of 3SG and 3PL  
39  
40 morphemes, using a similar picture selection task as in the above-described studies on  
41  
42 English, Spanish, German, and French. Participants in our study were two- and three-year-old  
43  
44 Dutch-speaking children. We predict that, due to the high phonological salience and strong  
45  
46 cue reliability of the Dutch 3PL morpheme, Dutch children will show earlier comprehension  
47  
48 of this morpheme than of 3SG. Moreover, given that 3PL has a higher salience and cue  
49  
50 reliability as compared to agreement morphemes in English, Spanish, and German, but not  
51  
52  
53  
54

---

55  
56 <sup>1</sup> In Polišenská (2010), a head-turn preference was also conducted among Dutch 18- and 19-month-olds,  
57 showing that Dutch infants were sensitive to 3SG agreement errors but not 3PL errors, as suggested by increased  
58 listening times. Different task materials were used than in the production task, however. Also, infants' looking  
59 behavior is indicative of early sensitivity to (un)grammaticality, and not necessarily comprehension.  
60

1  
2  
3 French, we predict that comprehension in Dutch is earlier than in English, Spanish and  
4  
5 German, but not necessarily earlier than in French. The second aim of our study is to compare  
6  
7 comprehension and production in the same children, using the same stimuli. In so doing, we  
8  
9 try to minimize the possibility of finding comprehension-production asymmetries that are due  
10  
11 to diverging methodologies (cf. Brandt-Kobebe & Höhle, 2010). We consider three possible  
12  
13 outcomes: that comprehension follows production, that production follows comprehension, or  
14  
15 that the two appear to be roughly contemporaneous.  
16  
17

## 18 Method

### 20 Participants

21  
22 Participants were 29 two-year-old Dutch monolingual children with a mean age of 31  
23  
24 months ( $SD = 3$ , range = 24 – 35) and 41 three-year-old Dutch monolingual children with a  
25  
26 mean age of 42 months ( $SD = 3$ , range = 36 – 47). These children were selected from a larger  
27  
28 sample of 103 children. Children were excluded from this sample if they suffered from  
29  
30 hearing impairment (1), had an elevated familial risk at language impairment or dyslexia (7),  
31  
32 were not monolingual Dutch (2), or if they had not completed the whole comprehension task  
33  
34 or at least half of the production task (20).<sup>2</sup> The group of two-year-olds contained 16 boys and  
35  
36 13 girls. The group of three-year-olds contained 19 boys and 22 girls. In both age groups, the  
37  
38 majority of children came from high SES families, defined as families in which one or both  
39  
40 parents had completed higher education (60% of two-year-olds; 64% of three-year-olds).  
41  
42 Children were recruited from four different geographic areas in the Netherlands. There is  
43  
44 much variation in verbal inflection in Dutch dialects (Bennis & MacLean, 2006), but this was  
45  
46 not an important factor in the current study, since all children were learning standard Dutch.  
47  
48  
49  
50

### 51 Materials

52  
53  
54  
55  
56  
57 <sup>2</sup> The inclusion criterion in the comprehension task was set at 100%, as there was a fairly high chance level in  
58 this task. In the production task, there was no such high chance level, so 50% completion was used as the  
59 criterion in this task to retain sufficient numbers for the analyses.  
60

1  
2  
3       **Comprehension task.** Comprehension was tested through a picture selection task in  
4  
5 which children selected one out of two pictures after listening to a sentence. The experiment  
6  
7 was modeled after the experiment reported in Brandt-Kobele and Höhle (2010) in that all  
8  
9 experimental items contained the Dutch pronoun *ze* ‘she/they’ that is ambiguous between a  
10  
11 (feminine) 3SG and 3PL meaning.<sup>3</sup> As such, the agreement morpheme on the verb provided  
12  
13 the only cue for disambiguating between a singular and plural meaning of the sentences (see  
14  
15 Figure 1 for an example item). The task was presented on a laptop and the presentation of the  
16  
17 stimuli was controlled through the experimental software E-prime 2.0 (Psychology Software  
18  
19 Tools, Pittsburgh, PA). Sentences had been prerecorded by a female speech therapist in a  
20  
21 child-friendly voice. The pictures were all high quality photographs of real situations, acted  
22  
23 out with playmobil figures. Recall that the pronoun ‘ze’ is ambiguous between third person  
24  
25 singular female and third person plural and thus, to ensure ambiguity of the subject pronoun,  
26  
27 all characters had female features.  
28  
29  
30  
31  
32  
33

34 [Insert Figure 1 about here]  
35  
36  
37

38       The task contained twelve experimental items, divided over a plural and a singular  
39  
40 condition, and six filler items. The verbs used in the experimental items were the following:  
41  
42 *aaien* ‘stroke’, *bouwen* ‘build’, *dragen* ‘carry’, *duwen* ‘push’, *gooien* ‘throw’, *kammen*  
43  
44 ‘comb’, *kijken* ‘watch’, *klimmen* ‘climb’, *lezen* ‘read’, *lopen* ‘walk’, *zoeken* ‘look for’, and  
45  
46 *slapen* ‘sleep’. Previous research on Dutch-speaking children has shown effects of coda  
47  
48 consonants on 3SG production such that ‘-t’ is produced more accurately on verbs of which  
49  
50 the stem ends in a sonorant rather than non-sonorant (Blom, Vasić, & De Jong, 2014).  
51  
52  
53  
54

55  
56 <sup>3</sup> Unlike in (colloquial) English, where the use of ‘they’ as a marker of 3SG (next to 3PL) is a recent  
57  
58 development, Dutch ‘ze’ (or ‘zij’) has been used as a marker of both 3SG and 3PL for a very long time. It was  
59  
60 used as marker of 3SG only in the extinct Germanic language Gothic, in the 6<sup>th</sup> century, but came into use as a  
marker of both 3SG and 3PL shortly thereafter, in the 7<sup>th</sup> or 8<sup>th</sup> century (van Bree, 1987).

1  
2  
3 Therefore, in our study, we controlled for properties of the coda: Half of the verbs ended in a  
4  
5 sonorant segment (i.e., nasal or vowel), the other half ended in a non-sonorant segment (i.e.,  
6  
7 obstruent or fricative). None of the verbs had consonant clusters in the coda of the verb stem,  
8  
9 to avoid phonological problems with consonant clusters in the 3SG condition. All verbs had  
10  
11 monosyllabic stems and were regular, high-frequency verbs that were selected from two  
12  
13 parental vocabulary checklists: the Dutch version of the MacArthur Communicative  
14  
15 Development Inventory (Zink & Lejaegere, 2002) and Lexilijst Nederlands (Schlichting &  
16  
17 Spelberg, 2002). Moreover, they all had low and very comparable age of acquisition ratings  
18  
19 according to an age of acquisition frequency list based on 30.000 Dutch words ( $M = 5.03$ ,  $SD$   
20  
21  $= 0.74$ , range = 3.34 – 5.78) (Brybaert et al., in press). The filler items tested children's  
22  
23 understanding of lexical contrasts, that is, the meaning of prepositions (e.g., *Ze staat*  
24  
25 *voor/achter het hek* 'She is standing in front of/behind the fence) or verbs (e.g., *Ze staat/zit op*  
26  
27 *de kist* 'She is standing/sitting on the box). All items were approximately equal in length,  
28  
29 containing four or five words and five or six syllables. For a list of the items, see the  
30  
31 Appendix.  
32  
33  
34  
35

36 The task started with three practice trials. In the first practice trial, children were  
37  
38 presented with a lexical contrast, to familiarize them with the procedure. They then received  
39  
40 two practice trials presenting pairs of one-actor and two-actor pictures. The aim of these trials  
41  
42 was to draw children's attention to the crucial contrast between the two pictures. In order not  
43  
44 to provide them with the targeted structures, children were presented with simple, verbless  
45  
46 sentences such as 'one girl on a seesaw' with as the accommodating pictures a picture of one  
47  
48 girl sitting on a seesaw and a picture of two girls sitting on a seesaw. If children selected the  
49  
50 incorrect picture in these practice trials, they were given feedback such that they noticed and  
51  
52 understood that the only difference between the two pictures was a difference in the number  
53  
54 of actors performing an action.  
55  
56  
57  
58  
59  
60

1  
2  
3 The items were presented in pseudo-randomized lists in which no more than two items  
4  
5 of the same type followed one another. Lists were counterbalanced across participants, such  
6  
7 that no child heard the 3SG and 3PL version of an item. Also, to control for left/right response  
8  
9 bias, additional lists were created in which the presentation of the pictures (left vs. right) was  
10  
11 varied. Forward and backward versions of all lists were used to control for any effects of  
12  
13 presentation order such as fatigue.  
14  
15

16 **Production task.** Production of subject-verb agreement was assessed with an elicited  
17  
18 production task in which children were asked to complete sentences produced by the  
19  
20 experimenter. In this task, children were presented with pairs of two pictures that presented  
21  
22 the same singular/plural contrast as in the comprehension task. To elicit agreement from the  
23  
24 children, they were provided with a prompt sentence and asked to finish this sentence.  
25  
26 Specifically, the experimenter described what happened on the first picture of a pair and then  
27  
28 prompted children to describe the second picture. The following is an example of a prompt  
29  
30 sentence: ‘*Look! These girls stroke the horse* [points to first picture in Figure 2] *and this girl*  
31  
32 *...?* [points to second picture in Figure 2].<sup>4</sup>  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

[Insert Figure 2 about here]

43 Note that the second picture in this pair is the same as the item used in the  
44  
45 comprehension task. This was done to minimize differences between the two tasks in terms of  
46  
47 stimulus materials, both in terms of the verbs used and the picture materials. To make sure  
48  
49 that children’s performance in the sentence completion task was not affected by the  
50

51  
52 <sup>4</sup> The demonstrative *dit* is used in Dutch with singular neuter nouns, whereas *deze* is used with plural nouns  
53 (regardless of gender) and singular common nouns. Hence, the demonstrative pronoun could provide the child  
54 with a cue as to the singularity/plurality of the subject. However, due to the Dutch system, there was no way to  
55 avoid having some marking of number on the subject DP. For instance, in the case of a definite determiner,  
56 *meisje* ‘girl’ is preceded by *het* ‘the’ that marks singular, neuter nouns, while *meisjes* ‘girls’ is preceded by *de*  
57 ‘the’ that marks plural nouns (regardless of gender). In the case of an indefinite determiner, *meisje* ‘girl’ is  
58 preceded by *een* ‘a’ that marks singularity while the plural indefinite equivalent would be the bare noun *meisjes*  
59 ‘girls’.  
60

1  
2  
3 comprehension task, the items were counterbalanced across the two tasks such that children  
4  
5 did not receive the exact same item in both tasks. Specifically, if they had been presented with  
6  
7 a 3SG sentence of a given item in the comprehension task, they were presented with the 3PL  
8  
9 sentence of that item in the production task.  
10

11  
12 As illustrated in Figure 2, the pictures depict an action being performed on different  
13  
14 objects (i.e., a horse vs. a dog). The reason for having a contrast in objects between the two  
15  
16 pictures of a pair was twofold. First, we expected the children to use sentences containing  
17  
18 verbs and objects instead of elliptic utterances (e.g., ‘This girl too’) if they had to describe a  
19  
20 contrast between objects. Second, elicitation of an object was considered important given that  
21  
22 Dutch children are known to go through a stage in which they produce infinitival clauses  
23  
24 (‘root infinitives’) instead of full-fledged finite sentences, such as *Mama koekje eten*  
25  
26 ‘Mommy cookie eat’ rather than *Mama eet [een] koekje* ‘Mommy eats a cookie’ (Jordens,  
27  
28 1990; Wijnen & Verrips, 1998). Crucially, use of a root infinitive is not considered to be an  
29  
30 agreement error (Poeppel & Wexler, 1993). In Dutch, the infinitive is homophonous with the  
31  
32 plural as it ends in ‘-en’. One way to distinguish between finite forms and root infinitives is to  
33  
34 determine a verb’s position relative to the object: whereas finite forms precede the object,  
35  
36 infinitives follow the object in Dutch. So, eliciting objects from the children was considered  
37  
38 important to be able to tell apart plural and infinitival verbs.  
39  
40  
41  
42

43  
44 The production task contained the same twelve experimental items and six filler items  
45  
46 as the comprehension task. It started with two practice trials to familiarize children with the  
47  
48 procedure. Experimental items were presented in counterbalanced lists in which no more than  
49  
50 two items of the same type were presented after one another. Of each list, forward and  
51  
52 backward lists were created to avoid order effects.  
53

#### 54 **Procedure**

55  
56  
57  
58  
59  
60

1  
2  
3 Children were tested individually by trained assistants in a quiet room at their day care  
4 centers or at home. In both tasks, assistants provided children with one more trial if children  
5 did not respond to the first one. If children did not respond to the second trial either, the item  
6 was scored as 'no response'. The comprehension task always preceded the production task.  
7  
8 Although this may have biased children's performance in favor of production, we decided not  
9 to counterbalance the order of the tasks, but always start with the less challenging  
10 comprehension task. Since, in the production task, a more active role is asked from the child,  
11 we feared that starting with this task would make the two- and three-year-old children shy or  
12 unwilling to participate. To reduce any transfer effects from the comprehension to the  
13 production task, two very different tasks (i.e., the Dutch Peabody Picture Vocabulary Test and  
14 a nonword repetition test) intervened between the two tasks. In addition, the items of the  
15 comprehension and production task were counterbalanced to make sure that children were not  
16 presented with the exact same items, as explained above. The entire session lasted  
17 approximately 30 minutes. At the end of the session, children received a small gift.

### 33 34 **Coding and Analyses**

35  
36 Children's responses in the comprehension task were scored online. Responses could  
37 be correct, incorrect, or null responses (2% of all responses). The responses in the production  
38 task were transcribed on the basis of video recordings, and coded for (i) verb type (target,  
39 semantically related, semantically unrelated, auxiliary, none), (ii) whether there was an object  
40 or not, (iii) the order in which the verb and the object appeared (target-like, or non-target-like,  
41 e.g., root infinitives), and (iv), agreement marking on the verb ('-t', '-en', null, other, e.g.,  
42 past participles). Scores were calculated as the percentage of correct responses out of the total  
43 number of correct and incorrect responses for each child. For production, scores were based  
44 on utterances that contained a target or semantically related verb and an object in target-like  
45 order only, that is, with the verb in V2 position.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 Repeated-measures ANOVAs were run on children's scores on the comprehension and  
4 production tasks to test for effects of condition (singular vs. plural) and group (two- vs. three-  
5 year-olds). Also, one sample t-tests with 50% correct as the reference level were run on the  
6 data of the comprehension task to see whether scores were significantly above chance.  
7  
8

9  
10  
11 A large number of responses in the production task could not be analyzed for subject-  
12 verb agreement, as they did not contain a verb at all or did not contain a verb in V2 position.  
13  
14 To avoid calculating percentages correct on the basis of few responses, only children who  
15 produced at least three analyzable (V2 utterances) per condition were taken into account in  
16 the ANOVA. As this reduced sample size (and hence the power of the ANOVA), in a second  
17 step, an additional logistic mixed-effect regression analysis was run. In this analysis, group  
18 and condition were entered as predictor variables and children's production accuracy on each  
19 item was entered as the (categorical) dependent variable (i.e., incorrect vs. correct). Subjects  
20 and items were entered as random factors. The aim of this analysis was to see if differences  
21 between groups and/or conditions were significant in an analysis on the basis of all data,  
22 rather than a small subset of the data, as logistic mixed-effect regression can take into account  
23 multiple responses of participants, including those with few responses.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37

38 Finally, to explore the relationship between comprehension and production, a Pearson  
39 correlation between children's comprehension and production scores was calculated.  
40  
41 Subsequently, individual children's scores on both tasks were inspected to see how many  
42 children performed better on production than on comprehension, how many children  
43 performed vice versa, and how many performed alike on both tasks.  
44  
45  
46  
47  
48

## 49 Results

### 50 Comprehension

51  
52 Mean accuracy scores (in percentages correct) and standard deviations on the  
53 comprehension task are presented in Table 2 for the two age groups separately.  
54  
55  
56  
57  
58  
59  
60



[Insert Table 2 about here]

A repeated-measures ANOVA with condition (singular vs. plural) as the within-subjects factor and group (two- vs. three-year-olds) as the between-subjects factor showed a main effect of condition ( $F(1,68) = 15.19, p < .001, \eta^2_p = .18$ ) and a main effect of group ( $F(1,68) = 5.98, p < .05, \eta^2_p = .08$ ). No interaction effect was found ( $p > .1$ ). These results indicate that both groups performed better on the plural than singular, with the three-year-olds outperforming the two-year-olds in both conditions. A t-test comparing performance against the 50%-chance level showed that, in the singular condition, performance was not above chance in either of the two groups ( $t(1,28) = -1.44, p > .05$ , Cohen's  $d = 0.54$  for the two-year-olds;  $t(1,40) = .11, p > .1$ , Cohen's  $d = 0.03$  for the three-year-olds). In the plural condition, however, both groups performed significantly above chance, with the effect being larger for the three-year-olds ( $t(1,40) = 5.34, p < .001$ , Cohen's  $d = 1.67$ ) than for the two-year-olds ( $t(1,28) = 2.39, p < .05$ , Cohen's  $d = 0.90$ ).

### Production

An overview of the types of responses given in the elicited production task is presented in Table 3.

[Insert Table 3 about here]

In both groups, most responses contained an object and a verb in target-like order (i.e., verb in V2 position) that could be analyzed for subject-verb agreement (69.1% for the two-year-olds; 89.5% for the three-year-olds). In the youngest age group, incomplete responses without a verb or with a verb in non-target position were also relatively frequent.

1  
2  
3 To avoid calculating accuracy scores for children on the basis of few data points,  
4 scores were calculated for children with at least three analyzable responses per condition.  
5  
6 Accuracy scores (in percentages correct) are presented in Table 4 for the two groups  
7  
8 separately.  
9  
10

11  
12  
13  
14 [Insert Table 4 about here]  
15  
16  
17

18 A repeated-measures ANOVA with condition as within-subjects factor and group as  
19 between-subjects factor showed no effects of condition or group and no interactions (all  $ps >$   
20 .1). Effect sizes were  $\eta^2_p = .03$  for both factors, however, suggesting that the lack of  
21 significant results may be due to a lack of power, as p-values are sensitive to sample size, but  
22 effect size is not. As can be seen in Table 4, there was a fairly small sample of two-year-olds,  
23 which may have led to insufficient power in the analysis  
24  
25  
26  
27  
28  
29  
30

31 To examine whether the above pattern of results would hold if more children were  
32 taken into account, an additional analysis was done in which responses from all children who  
33 produced at least one response with a verb in V2 position were taken into account, rather than  
34 only those responses from children who produced a minimal number of three analyzable  
35 utterances per condition. Specifically, we performed a linear logistic mixed-effect regression  
36 analysis with children's accuracy on each item as the dependent variable (0 = incorrect, 1 =  
37 correct), and group and condition as the predictor variables. Items and participants were  
38 entered as random factors in this analysis (see Method section). Mean accuracy scores for the  
39 two groups based on this larger sample of children are shown in Table 5. Table 6 presents the  
40 results of the logistic regression analysis.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

53  
54  
55  
56 [Insert Table 5 about here]  
57  
58  
59  
60

[Insert Table 6 about here]

Both group and condition have marginally significant effects on children's production accuracy, such that there is a trend for (i) the three-year-olds to outperform the two-year-olds, and (ii) performance to be better on the plural than on the singular. These results match the data in Table 4 above for the smaller group that showed numerically (but not statistically) higher accuracy for the three- than two-year-olds and for the plural than for the singular.

Overall, the production data in Table 5 show much lower accuracy than in the elicited production study on Dutch-speaking three-year-olds by Polišenská (2010) (see Introduction). One possible reason for this discrepancy may be a difference in design across studies. Whereas Polišenská did not provide children with a target verb in the prompt sentence, the current participants were provided with a verb form by the experimenter and then had to produce the same verb but in a different form in their own responses. Possibly, this task of producing the same verb as the experimenter but in a different form made the task more challenging for the children. If so, one would expect the current participants to often repeat the verb form produced by the experimenter, rather than leave out the inflection morpheme, a common agreement error in Dutch children (Blom, 2003; Blom, Polišenská, & Weerman, 2006).

The data in Table 7 suggest that this is indeed what happened. This table shows the frequency of the types of agreement errors made in both conditions for all children who produced at least one error involving an incorrect agreement morpheme (20 two-year-olds and 31 three-year-olds). In the singular condition, agreement errors involved verbs ending in '-en' more often than verbs ending in a 0-morpheme in both the two-year-olds (38% vs. 62%) and three-year-olds (40% vs. 60%). In the plural condition, agreement errors involved verb forms

1  
2  
3 ending in '-t' more often than forms ending in -0, both in the two-year-olds (27% vs. 73%)  
4  
5 and three-year-olds (16% vs. 84%).  
6  
7  
8

9  
10 [Insert Table 7 about here]  
11  
12

### 13 14 **Relation between comprehension and production** 15

16 To investigate how children's performance on the two tasks was interrelated, we  
17 correlated children's accuracy scores on both tasks. As in the previous analyses, production  
18 scores were based on children's agreement marking on analyzable structures only, and  
19 children had to produce at least three V2 structures per condition in order to be included in the  
20 analysis. A moderate, positive significant relationship between comprehension and production  
21 was found ( $r(52) = .38, p < .01$ ). An anonymous reviewer pointed out that this analysis may  
22 be problematic as children did not produce analyzable utterances for all verbs in the  
23 production task, hence different verbs may be compared across tasks. Therefore, we repeated  
24 our analysis for the children who responded to at least ten items in the production task (as a  
25 stricter criterion of all twelve utterances would result in a too small sample of only six  
26 children). The correlation for this sample of children with (nearly) complete overlap in verbs  
27 responded to across tasks, yielded a positive and significant correlation between  
28 comprehension and production that was slightly larger than found above ( $r(31) = .43, p = .01$ ).  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

45 Closer inspection of the data showed that of all 52 children, 30 children performed at  
46 least 10% less accurately on the comprehension task than on the production task (58%), 11  
47 children scored more or less the same on both tasks (i.e., less than 10% difference in scores)  
48 (21%), and 11 children performed at least 10% more accurately on the comprehension task  
49 than on the production task (21%). No differences in distribution were found between the two  
50 age groups.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 In summary, the present results showed that Dutch two- and three-year old children  
4 performed significantly above chance on 3PL, but not on 3SG, in a picture selection task.  
5  
6 Elicited production data from the same children showed that both age groups obtained  
7 accuracy scores between 64% and 78% on 3SG and 3PL. Analyses on the full set of  
8  
9 analyzable responses showed, moreover, that there was a trend for both groups to perform  
10  
11 better on the plural than on the singular. Accuracy in production was relatively low, however,  
12  
13 as compared with earlier studies. This may be due to children repeating the verb form  
14  
15 provided by the experimenter (as a prime) relatively often, as suggested by a comparison of  
16  
17 error rates: in both conditions, children used the form produced in the prompt sentence more  
18  
19 often than a null form, not produced by the experimenter. Finally, a moderate, significant  
20  
21 correlation ( $r = .38$ ) between children's comprehension and production performance was  
22  
23 found. Closer examination of the individual data showed that over half of the children (58%)  
24  
25 scored better on comprehension than on production, whereas the remaining children scored  
26  
27 better on production, or performed more or less alike on both tasks.  
28  
29  
30  
31  
32

### 33 Discussion

34  
35  
36 The results of this study indicate that the comprehension of verbal number agreement  
37  
38 in Dutch emerges at two years, but follows an asymmetrical pattern, as children this young  
39  
40 show comprehension of the 3PL, but not the 3SG morpheme. At three years, performance is  
41  
42 significantly higher, but shows the same pattern, suggesting that the comprehension of 3SG  
43  
44 develops more slowly. These findings differ from those on English-, Spanish-, German- and  
45  
46 Xhosa-speaking children who do not show above-chance performance on either 3SG or 3PL  
47  
48 before the age of five in very similar comprehension tasks. They also differ from findings for  
49  
50 French showing that, as early as at 2;6 years, French-speaking children show comprehension  
51  
52 of a special case of number agreement (liaison) of both 3SG and 3PL. So, in terms of the rate  
53  
54 of acquisition of verbal number agreement, Dutch is in between languages such as Spanish,  
55  
56  
57  
58  
59  
60

1  
2  
3 English, and German that mark the singular and/or plural with monosyllabic suffixes such as  
4  
5 /s/, /n/ or /t/, and French that incidentally marks plural with the special liaison morpheme /z/.

6  
7 The finding that Dutch children's comprehension of '-en' (but not '-t') emerges as  
8  
9 early as at two years supports the idea that morphemes that are perceptually salient and have a  
10  
11 high cue reliability are comprehended earlier than morphemes that are less salient and provide  
12  
13 less reliable cues (Brown, 1973; Legendre et al., 2013; MacWhinney et al., 1984). On the  
14  
15 basis of these data, we cannot tell apart, however, which one of these factors is more  
16  
17 important given that Dutch '-en' is both perceptually salient and provides a more reliable cue  
18  
19 as compared to '-t'. Another factor that may facilitate the acquisition of 3PL as compared to  
20  
21 3SG is that it involves syllabic rather than consonantal inflection, adding an extra syllable to  
22  
23 the verb stem. Previous studies have shown that syllabic morphemes are acquired before non-  
24  
25 syllabic ones by children with SLI (Leonard & Bertolini, 1998). Also, input frequency of both  
26  
27 specific morphemes and verb forms may play a role. In the current study, frequency was not  
28  
29 manipulated as a factor and thus could not be investigated. Future research could examine  
30  
31 how factors such as phonological salience, cue reliability, syllabicity, and input frequency  
32  
33 (type and token frequency) relate to, and possibly interact, in children's comprehension and  
34  
35 production of agreement.  
36  
37  
38  
39

40 One of the goals of the current study was to see whether there would be an asymmetry  
41  
42 between comprehension and production when the methodology to examine comprehension  
43  
44 and production was kept as constant as possible. Our results show that such an asymmetry  
45  
46 still exists, with production being more advanced than comprehension, but crucially, that it is  
47  
48 much smaller than assumed in previous research. The relatively low production rates in the  
49  
50 current study as compared to spontaneous speech are in line with results for English that  
51  
52 showing English-speaking children reach 90% accuracy in elicited production only by age  
53  
54 four (Rice & Wexler, 2002), about one to two years later than in spontaneous speech (Brown,  
55  
56  
57  
58  
59  
60

1  
2  
3 1973). They also align with previous findings for German showing that 3SG and 3PL  
4  
5 production rates, assessed through parental report, were much lower than reported for  
6  
7 spontaneous speech (Brandt-Kobele & Höhle, 2010). This suggests that specific task demands  
8  
9 play an important role, an observation that was further strengthened by the discrepancy  
10  
11 between the current results and those of Polišenská (2010) who found much higher accuracy  
12  
13 rates for Dutch three-year-olds, a difference that seems at least in part due to the prompt  
14  
15 sentences used.  
16

17  
18 A drawback of studies looking into the relation between comprehension and  
19  
20 production is that direct comparisons between the two domains are difficult, even if, as in the  
21  
22 current study, the exact same materials are used to assess both skills. It seems that strong  
23  
24 conclusions about a possible gap are only warranted if performance is either very low or very  
25  
26 high, such that there is a clear contrast. In the current study, there was a slight above-chance  
27  
28 performance in comprehension coupled with 60-70% accuracy in production. What do results  
29  
30 like these tell us about a possible comprehension-production asymmetry? This is a difficult  
31  
32 question that may lack a clear answer. Yet, we believe that the current results can contribute  
33  
34 to a better understanding of the relation between production and comprehension of subject-  
35  
36 verb agreement in Dutch in three ways. First, both in comprehension and production,  
37  
38 performance on 3PL was more advanced than on 3SG (either statistically or numerically),  
39  
40 which suggests that perceptual salience and/or cue reliability are important for both domains.  
41  
42 Second, comprehension and production were only moderately correlated, suggesting parallel  
43  
44 development in both domains, but only to some degree. Finally, we found that just over half  
45  
46 of the children (58%) obtained higher scores on comprehension than on production, but the  
47  
48 other half did not, suggesting considerable individual variation in children's relative  
49  
50 performance on both tasks. Taken together, these results show that there are parallels between  
51  
52 comprehension and production of subject-verb agreement in Dutch, with production being  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 slightly more advanced than comprehension at the group level (but not necessarily so at the  
4  
5 individual level), and less clearly so than assumed in previous studies on other languages.  
6

7  
8 Comprehension-production asymmetries have also been found for other areas of  
9  
10 language acquisition, such as the production and comprehension of the German particle *auch*  
11  
12 ‘also’ (Höhle, et al., 2009; Hüttner, et al., 2004; Müller et al., 2009) and pronouns in English,  
13  
14 Dutch, French, and German (Hendriks & Spender, 2006 for English; Weissenborn, Kail, &  
15  
16 Friederici, 1990 for Dutch, French, and German). This raises important issues about the origin  
17  
18 of such asymmetries. Previous accounts have proposed that extra-linguistic factors play a role  
19  
20 such as limited cognitive factors (Grodzinsky & Reinhart, 1993) or pragmatic factors  
21  
22 (Thornton & Wexler, 1999), or explained such asymmetries within an Optimality Theory  
23  
24 framework (Hendriks, 2008). Others have argued that the asymmetries are due to  
25  
26 methodology, such as different demands of comprehension and production tasks (Hirsh-Pasek  
27  
28 & Golinkoff, 1996) or side-effects of the comprehension tasks used (Brandt-Kobebe & Höhle,  
29  
30 2010). The present results suggest that asymmetries are still found if task materials across the  
31  
32 two domains are kept as similar as possible and if comprehension and production are assessed  
33  
34 in the same children. However the results also show that there are clear individual differences.  
35  
36 In the current study, about half of the children performed better on comprehension than on  
37  
38 production (58%), but the remaining children did not. Also, even though the use of the  
39  
40 materials in both domains provides a relatively good test of comprehension-production  
41  
42 asymmetries, it does not preclude asymmetrical findings due to diverging methodological  
43  
44 demands: listening to a sentence and choosing one out of two minimally differing pictures is  
45  
46 likely to call on different processes than producing a sentence containing a correct agreement  
47  
48 morpheme. One explanation for the current results that does take into account individual  
49  
50 variation, then, is that picture selection tasks place high demands on children’s information  
51  
52 processing system, and that children vary in their information processing skills. In order to  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 perform this task, children have to listen to a sentence, keep this sentence in mind, and – at  
4  
5 the same time – visually process two pictures (as well as the difference between these), and  
6  
7 then make a motor response. A question left for further research is whether a comprehension-  
8  
9 production asymmetry would still be found if a less demanding comprehension task such as,  
10  
11 for example, a preferential looking task, was paired with a relatively demanding production  
12  
13 task such as the current one.

14  
15  
16         Conversely, one factor that could contribute to the late comprehension of verbal  
17  
18 number agreement – next to or instead of task demands – is its redundant character.  
19  
20 Singularity/plurality is typically also encoded on the subject in non-pro-drop languages such  
21  
22 as Dutch, so children may be able to convey a singular or plural meaning without fully  
23  
24 grasping the meaning expressed by agreement morphemes (provided they know how to use  
25  
26 plural morphology on nouns). On this account, the inflected verb forms they produce would  
27  
28 be taken over directly from the input without being fully analyzed, as part of larger structures  
29  
30 (Theakston, Lieven, & Tomasello, 2003), and be comprehended only later in development.  
31  
32 Studies using nonsense verbs (Polišenská, 2010) or detailed analyses on productivity of  
33  
34 morphemes (Gathercole et al., 2002) suggest, however, that subject-verb agreement marking  
35  
36 may be used productively from early on, at least in some children, which is hard to reconcile  
37  
38 with this idea.  
39  
40  
41

42  
43         Before we conclude that Dutch children's comprehension of the 3PL morpheme  
44  
45 emerges as young as at age two, it is worth considering two alternative interpretations of our  
46  
47 results. According to the first interpretation, children would show above-chance performance  
48  
49 on 3PL but not 3SG because of an overall preference for 3PL pictures, regardless of  
50  
51 condition. Recall from above that Brandt-Kobele and Höhle (2010) found that German three-  
52  
53 to four-year-olds had a preference for plural pictures in a preferential looking experiment  
54  
55 containing the same type of sentences as studied here, also in the baseline phase of the  
56  
57  
58  
59  
60

1  
2  
3 experiment (i.e., prior to the presentation of the sentences). To explain this preference, they  
4  
5 propose that the two-actor pictures attract more attention from young children because they  
6  
7 are informationally more complex. The current results are compatible with such an overall  
8  
9 preference, as children selected the plural pictures more often than the singular pictures in  
10  
11 both conditions. However, it does not seem likely that our results are due to an overall  
12  
13 preference for two-actor pictures for two reasons. First, Brandt-Kobebe and Höhle found that  
14  
15 the preference for two-actor pictures was only visible in eye movements and not in children's  
16  
17 pointing in a picture selection task. This makes it unlikely that Dutch children would show  
18  
19 such a preference in pointing. Second, it is not clear why such a tendency would be stronger  
20  
21 for three-year-olds than for two-year-olds, as found in the current study, because children's  
22  
23 information processing skills should increase rather than decrease with age. For these reasons,  
24  
25 we think it is safe to assume that our comprehension data reflect Dutch children's ability to  
26  
27 comprehend verbal number agreement rather than a non-linguistic overall preference for  
28  
29 informationally more complex pictures.  
30  
31  
32

33  
34 The second, alternative interpretation of our comprehension results would hold that  
35  
36 children opted for the plural pictures most because a singular sentence is also compatible with  
37  
38 two-actor pictures (i.e., for one out of the two actors depicted), but not vice versa. While we  
39  
40 cannot rule out this possibility on the basis of the present data, it does not seem likely either.  
41  
42 First, French-speaking children could match 3SG as well as 3PL-sentences to their correct pictures in a  
43  
44 similar pointing task involving a contrast between one and two actors already at 30 months of  
45  
46 age (Legendre et al., 2010). Second, in a preferential looking experiment, both French- and  
47  
48 German-speaking children looked longer to the matching pictures in both the 3SG and 3PL  
49  
50 conditions, revealing sensitivity to 3PL as well as 3SG (Legendre et al., 2010; Brandt-Kobebe  
51  
52 & Höhle, 2010). These results show that young children may associate 3SG-sentences with  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 one-actor pictures, making it unlikely that Dutch children of an older age would not be able to  
4  
5 do so.

6  
7 A question left for future research is when Dutch children will show full mastery of  
8  
9 verbal number agreement in comprehension and production. For comprehension, closer  
10  
11 inspection of our data showed that, in our sample, even children who were nearly four years  
12  
13 old often did not show above-chance comprehension of 3SG. In our total sample, only one  
14  
15 child showed 100% mastery on 3SG and 3PL in the comprehension task, and three children  
16  
17 showed near-mastery (over 90% correct on both conditions). For production, ten children  
18  
19 showed 100% mastery on 3SG and 3PL in the production task, and two children showed 90%  
20  
21 mastery on this task, all of them being three-year-olds. In the two-year-olds', responses often  
22  
23 did not involve the targeted (verb-object) constructions, but lacked a verb or an object or had  
24  
25 an incorrect order (i.e., root infinitives). While this posed a problem for our analyses in terms  
26  
27 of power, the two-year olds' frequent omission of (finite) verbs may be interesting in itself as  
28  
29 it could reveal avoidance of subject-verb agreement and reflect, at least in part, young  
30  
31 children's difficulties with producing subject-verb agreement. Future research could  
32  
33 investigate whether comprehension of the 3SG morpheme is earlier, later or follows the same  
34  
35 time course as in previously investigated languages such as English and German in which  
36  
37 morphemes are also phonologically not very salient and the same morpheme can be used to  
38  
39 mark several functions within the paradigm. Ideally, this research involves older children, at  
40  
41 least up until age five, to be able to track the acquisition process from beginning to end and  
42  
43 compare the results cross-linguistically across age groups.  
44  
45  
46  
47  
48

49  
50 Also, future research could use more advanced methods of studying comprehension  
51  
52 than the current study. As outlined above, earlier studies that coupled (manual) forced-choice  
53  
54 picture selection tasks with preferential looking tasks (Brandt-Kobele & Höhle, 2010;  
55  
56 Legendre et al., 2010) have shown that the latter type of tasks may reveal knowledge of  
57  
58  
59  
60

1  
2  
3 subject-verb agreement in young children that does not become apparent in children's  
4  
5 pointing behavior. In the current study, we opted for a picture selection task, because our aim  
6  
7 was to extend earlier studies on English, German, and French that all used picture selection  
8  
9 tasks to a new language. There are, however, a few drawbacks to using picture selection as a  
10  
11 measure of comprehension. One is that performance is compared against chance, and one may  
12  
13 wonder whether above-chance performance is a good indicator of comprehension. Another  
14  
15 limitation of the task is its demanding character: Children not only have to integrate auditory  
16  
17 and visual information but also have to make an explicit judgment about sentences in the form  
18  
19 of a motor response. This demanding task, together with its limited informativeness (i.e.,  
20  
21 comparing against chance level as a measure of comprehension), may make picture selection  
22  
23 less than ideal for studying comprehension in young children. Future research on  
24  
25 comprehension-production asymmetries should explore the use of other methods to assess  
26  
27 comprehension, such as eye tracking or neurocognitive methods, to obtain data that reflect  
28  
29 children's comprehension abilities more closely.  
30  
31  
32  
33

34 Taken together, the results of this study show that Dutch children aged as young as  
35  
36 two and three years show above-chance comprehension of the agreement morpheme marking  
37  
38 3PL, but not 3SG, supporting previous claims that phonological salience and cue reliability  
39  
40 foster comprehension. As suggested in previous work, there was an asymmetry between  
41  
42 comprehension and production, despite complete overlap in task stimuli across the two  
43  
44 domains. This asymmetry, however, was much smaller than assumed in previous studies on  
45  
46 the basis of spontaneous speech. This suggests that, apart from linguistic properties of the  
47  
48 target language, specific task demands have a clear impact on children's performance.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- Bennis, H., & MacLean, A. (2006). Variation in verbal inflection in Dutch dialects. *Morphology, 16*, 291–312.
- Blom, E. (2003). From root infinitive to finite sentence. The acquisition of verbal inflections and auxiliaries. Utrecht: LOT dissertation series.
- Blom, E., Polišenská, D. & Weerman, F. (2006). Effects of age on the acquisition of agreement inflection. *Morphology, 16*, 313–336.
- Blom, E., Polišenská, D., & Weerman, F. (2008). Determiners, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second Language Research, 24*, 289–323.
- Blom, E. & Wijnen, F. (2013). Optionality of finiteness: Evidence for a no overlap stage in Dutch child language. *First language, 33*, 225–245.
- Blom, E., Vasić, N., & de Jong, J. (2014). Production and processing of subject-verb agreement in monolingual Dutch children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, Published online 09-05-2014.
- Brandt-Kobele, O.-C., & Höhle, B. (2010). What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua, 120*, 1910–1925.
- van Bree, C. (1987). *Historische grammatica van het Nederlands* [Historical grammar of Dutch], Dordrecht: Foris.
- Brown, R. (1973). *A first language: The early stages*. Oxford: Harvard University Press.
- Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (in press). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*.
- Clahsen, H. (1986). Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics, 24*, 79–121.

1  
2  
3 Gathercole, V. C. M., Sebastián, E., & Soto, P. (2002). The emergence of linguistic person in  
4  
5 Spanish-speaking children. *Language Learning*, 52, 679–722.

6  
7 Grodzinsky, Y., & Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic*  
8  
9 *Inquiry*, 24, 69-101.

10  
11 Gxilishe, S., Smouse, M., Xhalisa, T., & de Villiers, J. G. (2009). Children's insensitivity to  
12  
13 information from the target of agreement: The case of Xhosa. In: Crawford, J., Otaki,  
14  
15 K., & Takahashi, M. (Eds.), *Proceedings of the 3rd Conference on Generative*  
16  
17 *Approaches to Language Acquisition North America* (pp. 46–53). Somerville, MA:  
18  
19 Cascadilla Press.

20  
21  
22 Hendriks, P. (2008). A unified explanation for production/comprehension asymmetries. In:  
23  
24 Anna Gavarró Algueró & M. João Freitas (Eds), *Proceedings of GALA 2007* (pp. 240–  
25  
26 251). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

27  
28  
29 Hendriks, P., & Spender, J. (2006). When production precedes comprehension: An  
30  
31 optimization approach to the acquisition of pronouns. *Language Acquisition*,  
32  
33 13, 319–349.

34  
35  
36 Hirsh-Pasek, K., & Golinkoff, R., (1996). *The origins of grammar: Evidence from early*  
37  
38 *language comprehension*. Cambridge: MA: MIT Press.

39  
40  
41 Höhle, B., Berger, F., Müller, A., Schmitz, M., & Weissenborn, J. (2009) Focus particles in  
42  
43 children's language. Production and comprehension of auch 'also' in German learners  
44  
45 from 1 year to 4 years of age. *Language Acquisition*, 16, 36–66.

46  
47  
48 Hüttner, T., Drenhaus, H., van de Vijver, R., & Weissenborn, J. (2004). The acquisition of the  
49  
50 German focus particle auch 'too': Comprehension does not always precede  
51  
52 production. In A. Brugos, L. Micciulla & C. E. Smith (Eds.), *Proceedings of the 28th*  
53  
54 *Annual Boston University Conference on Language Development*. Somerville, MA:  
55  
56 Cascadilla Press.

- 1  
2  
3 Johnson, V. A., de Villiers, J. G., & Seymour, H. N. (2005). Agreement without  
4  
5 understanding? The case of third person singular /s/. *First Language*, 25, 317–317.  
6  
7 Johnson, E. K. (2005). Grammatical gender and early word recognition in Dutch. In  
8  
9 Proceedings of the 29th Annual Boston University Conference on Language  
10  
11 Development. Somerville, MA: Cascadilla Press.  
12  
13  
14 Jordens, P. (1990). The acquisition of verb placement. *Linguistics*, 28, 1407–1448.  
15  
16 Legendre, G., Barrière, I., Goyet, L., & Nazzi, T. (2010). Comprehension of infrequent  
17  
18 subject-verb agreement forms: Evidence from French-learning children. *Child*  
19  
20 *Development*, 81, 1859–1875.  
21  
22  
23 Leonard, L., Bertolini, U., Caselli, M., McGregor, K. & Sabbadini, L. (1992). Morphological  
24  
25 deficits in children with specific language impairment: The status of features in the  
26  
27 underlying grammar. *Language Acquisition*, 2, 151-179.  
28  
29  
30 Leonard, L., & Bertolini, U. (1998). Grammatical morphology and the role of weak syllables  
31  
32 in the speech of Italian-speaking children with specific language impairment. *Journal*  
33  
34 *of Speech, Language, and Hearing Research*, 41, 1363–1374.  
35  
36  
37 MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in  
38  
39 English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23,  
40  
41 127–150.  
42  
43 Montrul, S.A. (2004). *The Acquisition of Spanish: Morphosyntactic Development in*  
44  
45 *Monolingual and Bilingual L1 Acquisition and Adult L2 Acquisition*. Amsterdam/  
46  
47 Philadelphia: John Benjamins.  
48  
49 Müller, A., Höhle, B., Schmitz, M., & Weissenborn, J. (2009). Information structural  
50  
51 constraints on children’s early language production: The acquisition of the focus  
52  
53 particle auch (‘also’) in German-learning 12- to 36-month-olds. *First Language*, 29,  
54  
55 373-399.  
56  
57  
58  
59  
60



- 1  
2  
3 Pérez-Leroux, A. T. (2005). Number problems in children. In Gurski, C. (Ed.), *Proceedings*  
4 *of the 2005 Canadian Linguistic Association Annual Conference*.  
5  
6 <http://westernlinguistics.ca/Publications/CLA-ACL/Perez-Leroux.pdf>.  
7  
8  
9  
10 Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early  
11 German. *Language*, 69, 1–33.  
12  
13  
14 Polišenská, D. (2010). Dutch children's acquisition of verbal and adjectival inflection.  
15  
16 Utrecht: LOT dissertation series.  
17  
18  
19 Rice, M., & Wexler, K. (1996). Toward tense as a clinical marker of specific language  
20 impairment in English-speaking children. *Journal of Speech, Language and Hearing*  
21 *Research*, 39, 1239–1257.  
22  
23  
24  
25 Rubino, R. B., & Pine, J. M. (1998). Subject–verb agreement in Brazilian Portuguese: What  
26 low error rates hide. *Journal of Child Language*, 25, 35–60.  
27  
28  
29  
30 Schlichting, L., & Spelberg, L. (2002). *Lexilijst Nederlands*. Pearson Assessment and  
31 Information B. V.  
32  
33  
34 Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2003). The role of the input in the  
35 acquisition of third person singular verbs in English. *Journal of Speech, Language and*  
36 *Hearing Research*, 46, 863–877.  
37  
38  
39  
40 Thornton, R., & Wexler, K. (1999). *Principle B, VP Ellipsis, and Interpretation*. Cambridge,  
41 MA: MIT Press.  
42  
43  
44  
45 De Villiers, J. G., & Gxilishe, S. (2008). The acquisition of number agreement in English and  
46 Xhosa. In Brucart, J. M., Gavarro, A., & Solà, J. (Eds.), *Merging Features:*  
47 *Computation, Interpretation, and Acquisition* (pp. 104–123). Oxford: Oxford  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 independent factors in acquisition? Children's comprehension of object pronouns in  
4  
5 Dutch, French and German. *First Language*, 10, 141-166.

7 Wijnen, F., & Verrips, M. (1998). The acquisition of Dutch syntax. In S. Gillis & A. De  
8  
9 Houwer (Eds.), *The Acquisition of Dutch* (pp. 233–299). Amsterdam: John Benjamins.

11 Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English.  
12  
13 *Journal of Child Language*, 30, 1–41.

15 Zink, I., & Lejaegere, M. (2002). *N-CDIs: Lijsten voor Communicatieve Ontwikkeling*.  
16  
17 *Aanpassing en hernormering van de MacArthur CDIs van Fenson et al.* [A CDI user's  
18  
19 manual with normative and validity data]. Leuven (Belgium)/Leusden  
20  
21 (the Netherlands): Acco.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Appendix

Experimental items of the comprehension and production tasks:

1. Ze aait/aaien het hondje 'She/they strokes/stroke the dog'
2. Ze bouwt/bouwen de toren 'She/they builds/build the tower'
3. Ze duwt/duwen de kar 'She/they pushes/push the cart'
4. Ze gooit/gooien de bal 'She/they throws/throw the ball'
5. Ze kamt/kammen de pop 'She/they combs/comb the doll'
6. Ze klimt/klimmen op de trap 'She/they climbs/climb the stair'
7. Ze kijkt/kijken naar de TV 'She/they watches/watch television'
8. Ze loopt/lopen op de straat 'She/they walks/walk in the street'
9. Ze slaapt/slappen in het bed 'She/they sleeps/sleep in the bed'
10. Ze zoekt/zoeken de fiets 'She/they looks for/look for the bike'
11. Ze draagt/dragen de mand 'She/they carries/carry the basket'
12. Ze leest/lezen het boek 'She/they reads/read the book'

Table 1

*Agreement Paradigm in Dutch (Present)*

	Dutch example	English translation
1SG	Ik zing- <b>Ø</b>	I sing
2SG	Jij zing- <b>t</b>	You sing
3SG	Hij/zij zing- <b>t</b>	He/she sings
1PL	Wij zing- <b>en</b>	We sing
2PL	Jullie zing- <b>en</b>	You sing
3PL	Zij zing- <b>en</b>	They sing

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 2

*Mean Scores and Standard Deviations on the Comprehension Task for the Two- and Three-Year-Olds (in %)*

	Two-year-olds			Three-year-olds		
	M	SD	N	M	SD	N
Singular	44.3	21.5	29	50.4	24.6	41
Plural	58.1	18.2	29	68.7	22.4	41

For Peer Review

Table 3

*Response Types in the Production Task (in % and (N))*

	Two-year-olds ( <i>N</i> = 29)	Three-year-olds ( <i>N</i> = 41)
Lexical verb, no object	12.7 (30)	2.4 (10)
Lexical verb, object (non-target order)	18.2 (43)	8.1 (34)
Lexical verb, object (target order)	69.1 (163)	89.5 (374)
Auxiliary verb	2.6 (9)	6.1 (30)
Other (no verb, unintelligible)	19.0 (66)	6.1 (30)
No response	10.6 (37)	2.8 (14)
Total number of responses	348	492

Table 4

*Mean Scores and Standard Deviations on the Production Task for the Two- and Three-Year-Olds (in %)*

	Two-year-olds			Three-year-olds		
	M	SD	N	M	SD	N
Singular	64.2	27.1	14	70.1	31.3	38
Plural	71.2	26.1	14	78.5	25.5	38

Table 5

*Correct and Incorrect Verb Forms in the Production Task per Condition for the Two- and Three-Year-Olds (in % and (N))*

	Two-year-olds (N = 25)		Three-year-olds (N = 41)	
	Correct	Incorrect	Correct	Incorrect
Singular	55.3 (42)	44.7 (34)	69.8 (134)	30.2 (58)
Plural	66.3 (59)	33.7 (30)	75.0 (135)	25.0 (45)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 6

*Model Estimates, Standard Error (SE), Z- and P-values of Mixed-Effect Regression*

*Model*

	Coefficient	SE	Z	p
(Intercept)	0.65	0.32	2.02	.044
Group	0.71	0.37	1.90	.057
Condition	-0.42	0.22	-1.89	.058

For Peer Review



Table 7

*Error Types per Condition in Production Task (in % and (N))*

	Two-year-olds ( <i>N</i> = 20)			Three-year-olds ( <i>N</i> = 31)		
	-0	-t (3SG)	-en (3PL)	-0	-t (3SG)	-en (3PL)
Singular	38.2 (13)	-	61.8 (21)	39.7 (23)	-	60.3 (35)
Plural	26.7 (8)	73.3 (22)	-	15.6 (7)	84.4 (38)	-

For Peer Review

Singular condition:

Plural condition:

Ze aait het hondje ('She strokes the dog')

Ze aaien het hondje ('They stroke the dog')



Figure 1.

Example item of the picture selection task

1  
2  
3  
4  
5  
6  
7 Singular condition:



Plural condition:



*Figure 2.*

Example item of the sentence-completion task

For Peer Review