Article

second language research

Second Language Research
29(2) 201–229

DOI: 10.1177/0267658312473471
slr.sagepub.com
SAGE

# Learning exponents of number on first exposure to an L2

**Susanne E. Carroll and Elizabeth (Nina) Widjaja**
University of Calgary, Canada

**Abstract**
Number lends itself to the study of how input interacts with transferred first language (L1) knowledge to facilitate or impede second language (L2) learning. We present data from adult English speakers exposed for the first time to Indonesian, a language that expresses number through bare noun phrases, reduplication and numeral + classifier constructions. Participants were trained on each construction and subsequently tested on both familiar expressions and novel ones. Learners acquired all three constructions. We discuss our results in terms of current theories of second language acquisition.

**Keywords**
Autonomous Induction Theory, bare noun phrases, classifiers, Indonesian, input, number, reduplication

## I Introduction

The principal goal of second language acquisition (SLA) research is to explain how learners process, represent, and produce a second language (L2). SLA theories assume that learners learn an L2 on the basis of what they already know but vary in their definitions of how first language (L1) knowledge ('transfer') constrains the initial-state and how it determines the course of L2 learning. Emergentist theories (Ellis, 1998; MacWhinney, 2005) make no sharp distinction between processing and knowledge representation (Clark, 1993: 14). Accordingly they predict massive transfer from the L1 because the patterns of the L1 are 'entrenched' in the sub-symbolic processing system that analyses L2 speech. For example, English expresses a singular/plural contrast instantiated on suffixes (*cat*/*cats, fish*/*fishes*) and uses bare noun phrases to express proper names (*John loves Mary*), SUBSTANCEs (*We ate fish for dinner*), plural THINGs

**Corresponding author:**
Susanne E Carroll, SS830, Department of Linguistics, University of Calgary, 2500 University Drive NW, Calgary, AB, T2N 1N4, Canada.
Email: susanne.carroll@ucalgary.ca

(*We ate potatoes with them*), or KINDS (*Fish have gills*).[1] Indonesian uses bare noun phrases (BARE) to refer to both single and multiple objects (*buku* '(a) book' and 'books'), but it also uses reduplication (REDUP) for plurals (*buku-buku* 'books') and a numeral + classifier construction or NUMCL (*lima buah buku* '5CL:books'). Emergentist theories predict that BARE will be processable but interpreted as plural THINGs, SUBSTANCEs or KINDs. The referential interpretation of BARE as a single THING, and the acquisition of REDUP and NUMCL constructions should emerge only slowly based on input. This is because the weights in the connectionist network will adjust to novel forms in the input only incrementally (MacWhinney, 2000; Williams, 2003). Since English has numeral + noun sequences (*three books*), we also expect Numerals + Nouns (NUMN) to emerge given NUMCL as input.

Schwartz and Sprouse (1994, 1996) predict massive transfer from the L1 for quite different reasons. Their 'Full Transfer/Full Access' (FT/FA) model stipulates that the L1 grammar is the initial state of the L2. Accordingly, BARE should be acquired early but exhibit only the meanings expressable in English. Since FT/FA trades in abstract morpho-syntactic constituents, e.g. Plural, as well as in algorithmic rules (e.g. Merge), if REDUP were to be acquired early on, this would be consistent with FT/FA since REDUP can be analysed as a base-suffix structure.[2] Since English does not have a functional category of classifiers, they should be more difficult to acquire.

The Minimal Trees Hypothesis or MTH (Vainikka and Young-Scholten, 1994, 1996) postulates that only lexical categories transfer from the L1. Functional categories are absent in the early stages of L2 acquisition. The MTH therefore predicts the early acquisition of BARE regardless of what occurs in the input and the patterns of the L1. Functional exponents of number should appear only subsequently and are predicted to emerge in conformity to a hypothesized universal functional hierarchy. Thus, if a Num(ber)P is motivated as the projection of number features (Ritter, 1991, 1992), MTH predicts that exponents of number (e.g. singular and plural), should emerge before higher order categories like DP (determiner phrase). In addition, if classifiers project a classifier phrase or CLP (Li and Shi, 2003: 23), and this CLP occurs between NumP and DP (the locus of deictics and specificity markers), classifiers should emerge after plurals.

Klein and Perdue (1997: 303) claim that early SLA is largely independent of the properties of the L1. Learners construct a 'Basic Variety' consisting of exemplars from lexical categories. The Basic Variety lacks functional categories and so may look like the reduced structures of Organic Grammar (Vainikka and Young-Scholten, 2006), but transfer plays no role in explaining why. Input properties and functional constraints determine the order in which functional categories emerge. Everything else being equal, BARE (with the appropriate interpretation) and NUMN should be acquired early on, but REDUP and NUMCL should be acquired only later.

Processability Theory (PT; Pienemann, 1998, 2007) claims that developmental stages are determined by constraints on speech planning and production. Specifically, the first stage will be limited to activating from memory the forms that are stored there (lemmas of simple words, irregular complex forms, and formulaic expressions). The next stage will involve producing productive morphology, such as the English regular plural marker. A subsequent stage will involve unifying constituents within a phrase, e.g. classifiers and a noun. Then phrases will be unified within a sentence. PT therefore predicts the early acquisition of lexical expressions and irregular functional categories, the subsequent acquisition of regular plurals and finally the acquisition of phrasal syntax. In fact, irregular plurals often do emerge early, and plural [s] is produced reliably on noun phrases before the English learner produces subject–verb agreement (Bailey et al., 1974; Dulay and Burt, 1973; Hakuta, 1976).

Finally, the Autonomous Induction Theory (AIT; Carroll, 2001) starts from empirical evidence showing transfer of L1 perceptual and processing procedures in order to argue that such L1 procedures must be the starting point for analysis of L2 input. Indeed, evidence for such transfer in later stages of acquisition is robust (see, amongst others, Archibald, 1993; Bohn and Flege, 1992; Dupoux et al., 1997; Jared and Kroll, 2001; Marian and Spivey, 2003; Weber and Cutler, 2004; Yamada and Tohkura, 1992). Studies of first exposure learners reveal that transferred processing procedures operate at that stage too. Initial-stage learners are more likely to segment L2 words if they resemble L1 words (Rast, 2008; Rast and Dommergue, 2003; see also Shoemaker and Rast, this issue), suggesting effects of L1 lexical organization on L2 segmentation. L2 speech can activate L1 lexical entries at the initial stage too (Carroll, 2012, in press).

In the course of language use, processors attune to the frequency with which forms and constructions are processed in the input or are produced in speech. Transfer reflects, therefore, not merely what is possible in a grammar (as FT/FA asserts), but also what is most probable given statistical properties of particular text types.[3] Like FT/FA, Organic Grammar, and PT, the AIT assumes that learners represent input in terms of abstract linguistic categories. Like PT, it assumes that morphemes and words unify into larger constituents. Unlike these other theories, it postulates that expressions are acquired as the result of complex interactions across autonomous levels of linguistic representation. In particular, it considers the relevance of prosodic and semantic information in providing 'routes' or mappings to the syntax. For such reasons, the AIT predicts that the sound forms of Bare, Redup and NumCl should be acquirable if the sound forms are consistent with prosodic transfer from English and the sound forms can be mapped to relevant conceptual structures. Once sound forms have been segmented and represented in memory, they will be mapped to conceptual structures and to morpho-syntactic structures if (1) meanings can be independently established from context, say, through inference from visual information, and (2) because of universal mappings from meaning to grammar.

To sum up, different SLA theories permit precise predictions about the acquisition of means to express number, based on their assumptions about transfer from the L1.

(1) a. Bare (NumN) < Redup/NumCl (Basic Variety, emergentist theories)

b. Bare (NumN)/Redup < NumCl (FT/FA)

c. Bare (NumN) < Redup < NumCl (MTH, PT)

d. Bare/Redup/NumCl (AIT)

# II Number in English and in Indonesian

## 1 Number in English

We assume that words occur in three autonomous representations, created in distinct modules of the language faculty (pace Jackendoff, 2002). In English, expressions like *a book, books* and *chicken* are represented in memory as in (2).[4]

```
(2) a. a book
      PWi          < -- >    NPi        < -- >  ⎛ BOOK          ⎞m
      / \                    / \                ⎜    /          ⎟
    Clij PWk               DETj  Nk             ⎜  INDEF        ⎟
     |    |                      |              ⎜ AMOUNT: 1     ⎟
     σ    σl               [+Count]             ⎝               ⎠i
     /   / | |                                   THING Imaginary
     ə   b ʊ k

    b. books
      PWi          < -- >    NPi        < -- >  ⎛ BOOK          ⎞
      /  \                   /  \               ⎜ AMOUNT: > 1   ⎟
    PWj  Afk                Nj   Plk            ⎝               ⎠i
    /| |   |                 |                   THING
    b ʊ k  s              [+ Count]

    c. chicken
      PWi          < -- >    NPi        < -- >  [CHICKEN ]i
      /  \                    |                  SUBSTANCE
      σ  σ                    N
      /| / | \                |
    tʃ I k ɨ n             [–Count]

   PW = Prosodic Word, Cli = Clitic, σ = Syllable
```

The leftmost diagrams in (2) are simplified representations of the sound form of the words. Diagrams in the middle express the word's morpho-syntactic properties. Diagrams on the right of (2) are conceptual structures. Correspondences across the levels are indicated by an index 'i', 'j', 'k', etc., which should be read as: Unit of Type A maps to Unit of Type B. In (2c), this index shows that the Prosodic Word (PW) maps to a noun phrase (NP) which, in turn, maps to a SUBSTANCE concept. We will start with discussion of conceptual structures, in particular, of the difference between reference and morpheme meaning (what is traditionally called sense).

Acquisition theories typically assume that visual and other non-linguistic cues are the basis for determining what a speaker is referring to (in ostensive reference) at early stages when a learner cannot rely on knowledge of the sense of expressions. In Conceptual Structure Semantics, the sense of lexical expressions is typically encoded in an ontological category (e.g. THING, PERSON, PATH, EVENT, and so on) that occurs in functor/argument structures (Jackendoff, 1983). In grasping the sense of *This is John* (while looking at a picture of John) the listener will represent JOHN as a PERSON category. In the sentence *These are books*, BOOK will be represented as a THING category. Reference is encoded formally in a 'referential tier' (Jackendoff, 2002: Chapter 10).[5] When a learner hears a sentence like *This is an apple*, while looking at an apple or at a picture of an apple, the consciously experienced APPLE will be encoded in a conceptual structure that bears an indexical feature ('m' in (2a)). Indexical features serve to bind descriptive features (for example, that a thing is approximately round, that it is red, that it is the size of an apple, etc.). The conceptual structure also permits a diacritic marking that a conceptual unit is external (the picture is of a real apple hanging from a real apple tree) or is imaginary (the picture is clearly an artist's rendition of an apple). Such features permit a valuation that can serve to ground beliefs that support truth value judgements, inferences to properties or to categories as well as motoric planning. KINDs are represented by ontological categories but lack such valuations or indices.

Reference does not require the representation of sense. In looking at apples or pictures of apples, we construct 'spatial structures', the locus for encoding perceptual information (Jackendoff, 2002: 12–14). Thus, a picture of an apple slice will be encoded in a spatial structure. The interpretation of the sentence *This is an apple slice* will be encoded in a conceptual structure as well.

As learners learn the *senses* of sound forms, they will construct ontological categories, functor–argument structures, semantic features, indices, and various other universal devices that Conceptual Structure Semantics makes use of. To date, studies of first exposure learners (Carroll, 2012, in press; Gullberg et al., 2010, in press) have shown that they can readily map referents to pictures on the basis of minimal input. On our story, that means that they construct spatial structures and can map the contents of a spatial structure to a sound form in the prosodic representation, via the referential tier of a conceptual structure. However, no work has been done so far to show that first exposure learners can acquire the senses of words from controlled input. This step is, however, critical since only the units of a conceptual structure will combine to build a proposition. Units of a spatial structure (constructed from processing a picture, say) do not combine to build the meanings of linguistic units.

Visual processing is not only a critical tool for language learners, but it can also shed light on the internal organization of linguistic cognition. For example, Barner and Snedeker (2005) have used pictures with adults to show that the mass/count distinction in English is a grammatical distinction (relevant to the middle diagrams in (2)), not a conceptual category. Mass nouns are typically assumed to refer to entities that cannot be individuated. Count nouns, as their name implies, can be counted. Barner and Snedeker (2005) showed English speakers a line drawing of two very large objects (or two large blobs of a substance) and asked them to compare that drawing to one of three much smaller exemplars of the same object (or three much smaller blobs of the same substance). They asked participants: *Who has more?* (*Who has more ketchup? Who has more shoes? Who has more furniture?*). By hypothesis, to answer the question, participants construct a spatial structure of the picture and link it to a conceptual structure of the question. They must then draw the correct inference, also encoded in a conceptual structure. Results indicate that English speakers count the number of THINGs (3 small shoes are 'more shoes' than 2 big ones) and assess the volume of SUBSTANCEs (2 large blobs of ketchup are 'more ketchup' than 3 little blobs). Crucially, they count AGGREGATEs (furniture, silverware, footwear, lingerie), which are grammatically mass. D MacDonald (2010) carried out a replication comparing native speakers of English to intermediate Korean L2 learners of English.[6] She showed that both groups quantified over number when processing count nouns, quantified over volume when processing bare mass nouns that refer to SUBSTANCEs (the *ketchup*-type mass nouns), and quantified over number when processing AGGREGATE mass nouns. Importantly, D MacDonald showed that the Koreans responded randomly to ambiguous cases where sensitivity to plural-marking is needed to determine the meaning: *Who has more stone?* / *Who has more stones?* Thus, on three classes of nouns, the Koreans appear to have mastered both the count/mass distinction and plural-marking. The ambiguous class shows that, in fact, they have not. By testing Koreans on comparable Korean questions, and comparing responses across the two languages by the semantic class of a noun, D MacDonald was able to show that her participants responded in the same way in Korean and English. Thus, native speakers of English were relying on linguistic knowledge of the English functional system, namely, the sense of plural-marking, in addition to the contents of conceptual structures and spatial structures to perform the task, but the Koreans appear to have relied solely on noun semantics and the contents of spatial structure representations. They were not making use of grammatical number.

In (2c), we have a representation of the noun *chicken* that might appear in a sentence like *I ate chicken for dinner*. This interpretation contrasts sharply with the THING-reading of *I ate a chicken for dinner*. In order to count out a SUBSTANCE, it must be measured out. Measurement phrases like *a grain of …, a drop of …, a cupful of …* and Indonesian classifiers can express these concepts.

As noted, English permits bare noun phrases denoting SUBSTANCEs and plurals (*I ate chicken for dinner* / *I ate chickens for dinner*). When Anglophones learn a language such as Indonesian where referents are typically realized as bare nouns, they must learn that bare nouns can express as well single tokens of THING concepts (as in (2a)). A transfer story predicts that learning this interpretation of Indonesian bare noun phrases will be difficult.

Let us now turn to sound structures. An early-stage learning problem involves segmenting novel sound forms from the speech signal to store in memory (Gullberg et al., 2010; Rast, 2008, 2010; Shoemaker and Rast, this issue). The Basic Variety Theory, FA/FT, Organic Grammar, and PT do not concern themselves with this aspect of language acquisition and consequently make no predictions. We think that it is a mistake to ignore potential interactions between the sound properties of linguistic units and their emergence as grammatical entities. The AIT predicts that sound form segmentation will depend on the prosodic properties of form in the input and transferred L1 segmentation biases (Carroll, 2004, 2010, to appear). Prosodic properties of the input will favour, in languages like English, the segmentation of Prosodic Words over the segmentation of Clitics and Affixes. In (2a), the expression [ə] is a prosodic Clitic, an element that attaches to a Prosodic Word but is not itself a Prosodic Word. In the target system, this Clitic maps to a Determiner category in the morpho-syntactic representation. This mapping is signalled by the use of an index, here ‘j’. In the target system, it maps to a semantic feature INDEF in the conceptual structure.[7] In (2b), the consonant /s/ appears as the prosodic unit Affix. This Affix unit maps to a Plural morpheme in the syntax, and maps to the meaning (approximately) ‘ > 1’ in the conceptual structure.

Some have argued that functional categories are missing in early stages due to objective properties of their sound forms in the input (Hatch, 1983; Larsen-Freeman, 1975). Ostensibly, functional categories are (universally) not perceptually salient because they are not stressed, their syllables are reduced, their vowels are centralized, and so on (Goldschneider and Dekeyser, 2001). Such properties are true of English Clitics and Affixes alike. The AIT claims, in contrast, that transferred parsing procedures are as important as objective properties of the signal. For example, an Anglophone should transfer a preference for Prosodic Words that start with a strong syllable (Cutler, 1988). Accordingly, such a learner should initially not represent any weak syllable to the left of a strong syllable when segmenting Prosodic Words but should represent a weak syllable to the right.[8] Thus, the AIT predicts that for such learners even functional categories should be learnable if they are instantiated in strong syllables or in weak syllables following strong syllables.[9] Since the sound form of Indonesian plurals is a copy of the bare noun and meets this requirement, segmentation should be easy.

Of course, segmenting sound forms is just a first step. The learner must map these sound forms to morpho-syntactic and conceptual correspondents. As seen above, English provides both a morpheme (Plural) and a meaning that can transfer. The AIT also relies on universal mappings to make predictions about L2 development. We assume that learners can draw on Universal Grammar to analyse reduplicated words as either base–suffix or prefix–base. In addition, in our study, learners were presented with pictures of objects. Each sound form was used referentially in context to talk about one or multiple objects. By hypothesis, learners construct spatial structures of what they have seen that are linked to elements in the referential tier of a conceptual structure. These elements, because they are referential, will be mapped to a Noun Phrase. The shared index *i* in (2) encodes the hypothesis that only phrases are referential in utterances, not words. Anglophones exposed to Indonesian bare noun phrases should analyse them as NPs. For the same reason, Anglophones should represent Redup and NumCl as NPs.

## 2 Number in Indonesian

Indonesian bare noun phrases refer to both single and multiple indefinite objects, as in (3):

(3) Saya mem-beli buku.[10]
I buy book
a. ‘I am buying a book.’
b. ‘I am buying books.’
c. * ‘I am buying the book.’
d. * ‘I am buying the books.’

To refer to objects mentioned previously in the discourse, a speaker will use a specificity marker *-nya* or a demonstrative *ini*/*itu* ‘this/that’.

**Table 1.** Indonesian reduplication.

| Bare noun | Gloss | Reduplication | Gloss |
|---|---|---|---|
| *Full reduplication*: | | | |
| *buku* | 'book' | *buku-buku* | 'books' |
| *anak* | 'child' | *anak-anak* | 'children' |
| *rumah* | 'house' | *rumah-rumah* | 'houses' |
| *Imitative reduplication*: | | | |
| *sayur* | 'vegetable' | *sayur-mayur* | 'vegetables' |
| *muda* | 'young' (adjective) | *muda-mudi* | 'youngsters' |
| – | | *teka-teki* | 'puzzle(s)' |

(4) a. Saya mem-beli bukunya
'I am buying this (particular) book.'
b. Saya mem-beli ini buku
'I am buying this book.'

Indonesian plurals are optional and involve reduplication.

(5) Saya mem-beli buku-buku.
'I am buying books.'

Example (5) illustrates full reduplication, the most frequent and productive form. Indonesian also has imitative reduplication where the reduplicant (the second half of the word) resembles the base with some change in vowels or consonants. A variant of this was used in the study.[11] See Table 1.

Indonesian classifiers are used to conceptually package matter as THINGs, or SUBSTANCEs (Aikhenvald, 2000; Allen, 1977). According to Widjaja (2010: 22), Indonesian classifiers provide information about the size, shape and animacy of referents. There are distinct classifiers for persons, animals and things:[12]

(6) a. dua (orang) murid
2Cl:human students
'two students'

b. lima (ekor) anjing
5Cl:animal dog
'five dogs'

c. tiga (buah) buku
3Cl:object book
'three books'

A classifier can express permanent or inherent properties of the referent of the noun, or temporary characteristics (Bisang, 1999). The classifiers in (6) express reasonably permanent properties of the referents. The examples in (7) show that a single object is conceptualized differently when occurring with two distinct classifiers *puntung* and *batang*. A *rokok* ‘cigarette’ is conceptualized as a long, thin entity, namely an intact (unsmoked) cigarette when classified by *batang*. When classified by *puntung* it is conceptualized as a short, truncated entity, viz. a cigarette butt.

(7) a. dua batang rokok
2Cl:long, thin cigarette
‘two cigarettes’

b. dua puntung rokok
2Cl:short, truncated cigarette
‘two cigarette butts’

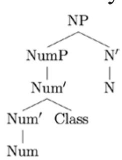
The examples in (8) illustrate how classifiers are used to individuate SUBSTANCEs.

(8) a. Dina me- nyiapkan se- cangkir teh untuk tamu -nya.
Dina ACT prepare 1Cl(cup)tea for guest.3rdPerson(Possessive)
‘Dina prepares a cup of tea for her guest.’

b. Saya punya dua helai kain cantik untuk kamu.
1stPerson have 2Cl(sheet)fabric beautiful for 2ndPerson
‘I have two beautiful pieces of fabric for you.’

R MacDonald (1976) and Chung (2000) claim that the Indonesian classifier system is disappearing, explaining why classifiers are optional and licensed only when numerals are also present (Bisang, 1999; Ernst, 1991). NumCl has the following structure (Dalrymple and Mofu, 2009).

(9) Numerally modified noun phrases




Let us now turn to the study.

## III The study

### 1 Participants

Data was gathered from 29 paid participants ranging in age from 18 to 38 ($M$ = 24.76, $SD$ = 7). All were native speakers of English.[13] None had any knowledge of Indonesian or any other Asian language.[14]

### 2 Methodology

The experiment was divided into two meetings with a two-week interval between sessions. Session 1 consisted of training and testing, and Session 2 involved re-testing. In the first meeting, participants had to correctly learn to name in Indonesian pictures of 20 objects via the three constructions. After successfully completing three learning tasks, participants performed a NoPrompt Elicitation task, in which they had to name from memory without prompting the items that they had previously learnt. Having successfully learnt to criterion in Session 1, participants were invited to return for Session 2 to perform two tasks: a retention task and a task measuring ability to generalize the rules of REDUP and NUMCL to 24 novel words (novel).

*a Procedures.* We trained all participants first on the nouns in the BARE task. This task was followed by either REDUP or NUMCL in counterbalanced order. The experiment was computer-operated, using standardized stimuli and instructions, and programmed using the E-prime platform. Instructions were presented in English orally using studio-quality headphones, and also visually on the computer screen.

*b Session 1: Learning trials to criterion, test, and NoPrompt Elicitation.* Each learning task consisted of training and verification trials (trial1, trial2, …, trial10), followed by the test phase. In the training trials, participants saw 20 pictures in random order and heard a sentence providing them with the name of the item that they were looking at. A verification phase followed in which the participants responded to a forced-choice question by pressing a key on the computer keyboard. They were to press F1 if the name of the item that they saw corresponded to the first word they heard or F12 if the name of the item corresponded to the second word they heard. Each question was separated from the next by 1,000 ms, during which time participants saw a fixation cross in the middle of the computer screen.[15] At the end of the verification stage, the participant's mean score was reported on the screen. If the participant did not reach criterion (100%) on a given training trial, the program automatically started a new training trial, consisting of presentation + verification phases. All stimuli during training trials were randomized. When successful, participants were tested in the same way as in the verification phase, with an additional production task. Answers were recorded on a digital audio recorder.

After completing the BARE task, participants learned REDUP and NUMCL in the same manner. After successfully completing all three learning tasks, participants performed the NoPrompt Elicitation task. Participants saw 20 pictures used in the first three tasks and, without prompting, were asked to name the depicted objects from memory. They saw eight pictures of a single item and 12 pictures of multiple items. Some of these were pictures of pieces or slices of objects since participants were trained on the NUMCL construction with such objects. See Appendix 1. Responses were recorded.

*c Session 2: Re-test and production of novel items.* Participants who reached criterion on all learning tasks were asked to return to the lab approximately two weeks later. Two tasks were conducted in the second meeting: a test similar to the test phase of the first meeting (re-test) to refresh their memory; and a production task of novel items (novel).

In re-test, a total of 40 test items from REDUP and NUMCL learning tasks were combined and divided into two test blocks. The same procedures from the test phase in Session 1 were followed. Participants also had to pronounce the names of the items. For the first 20 items, participants were given feedback immediately after each response.

Their mean score was also reported on the screen. For the next 20 items, participants did not receive feedback, but were provided the mean score at the end.

In the novel task participants were tested on their production of 24 novel items. Since the participants had not yet been exposed to these items, responses required that they draw on the induced rule for constructing reduplicated plurals or that they combine one of the classifiers with the novel word. Participants saw a picture on the screen, and heard the name in one of the three constructions they had been trained on. A fixation cross appeared for 1,000 ms followed by a different picture of the same item. Participants were asked to name the object using a form different from the form used in the prompt.[16] For example, they first heard the prompt in the bare form – *ini adalah tomat* ('this is (a) tomato') – coupled with a picture of one tomato. Next, they saw a picture of three individual tomatoes. Participants then had the option to produce either the reduplicated form *ini tomat-tomat* ('these are tomatoes'), or the numeral classifier form *ini ta pak tomat* (three CL. tomato; 'these are three tomatoes'). If they saw a picture of two apple slices, they were to produce NumCl creating occasions to produce NumN without the classifier. They were given 10,000 ms to respond, and pressed the space bar to hear the next item. Answers were recorded on a digital audio recorder.

*d Stimuli.* For each visually represented object, four pictures were created: a picture of a single item, and pictures of two items, three items and five items. For Bare, pictures of single and multiple items were used to show learners that bare nouns can represent single and multiple objects, e.g. a single playing card, or three lemons. For Redup, only pictures of multiple items were used, e.g. three bottles of water, or a pile of seven books. In NumCl, pictures of two and three objects (or two or three portions of objects) were shown, coupled with the numeral '2' or '3' typed above the pictures. See Appendix 1 for sample pictures.

Since the focus of the experiment was on learning the constructions rather than vocabulary, word learning was designed to be as simple as possible. For ease of pronunciation, the nouns chosen were bi-syllabic and consisted of mostly CV syllables. To further simplify word learning, half of the nouns were English cognates, since first exposure studies have shown that cognates are easier to learn than non-cognate words (Carroll, 2012). The list of items is provided in Table 2, organized by the classifier used (shown in parentheses).

Three versions of simple sentences and forced-choice questions were recorded for each test item by the second author (a native speaker of Indonesian). For each trial of the learning tasks, items were presented in the following frame: *ini adalah …* ('this is/these are …') where the dots were replaced by a target item. In the test and re-test phases, the presentation of each item was prompted with the frame: *apakah ini ..., atau ...?* ('is this/are these …, or …?'). In the questions, the target item appeared in the first or second NP of the conjunct. The alternative was chosen from the set of nouns presented. In the case of Redup, the forced choice alternative differed minimally from the target, e.g. target *kuda-kuda* (full reduplication) versus alternative *kudi-kuda* or target *leman-lemon* (imitative reduplication) versus alternative *limin-lemon*. In some cases, the alternative to an imitative reduplication (e.g. *sapa-sapi*) was a full reduplication (e.g. target *sapi-sapi*). In the case of NumCl, the forced choice alternative involved in some cases the other

**Table 2.** List of stimuli.

| Animal | | Object | | Slice | |
|---|---|---|---|---|---|
| *Items for trials, test, re-test and elicitation task*: | | | | | |
| **zebra** | 'zebra' | ***apel*** | 'apple' | ***apel*** | 'apple' |
| **panda** | 'panda' | ***lemon*** | 'lemon' | ***lemon*** | 'lemon' |
| ilkan | 'fish' | ***coklat*** | 'chocolate' | ***coklat*** | 'chocolate' |
| kuda | 'horse' | **kartu** | 'card' | pita | 'ribbon' |
| babi | 'pig' | **bola** | 'ball' | kue | 'cake' |
| ayam | 'chicken' | **koin** | 'coin' | roti | 'bread' |
| bebek | 'duck' | **balon** | 'baloon' | keju | 'cheese' |
| sapi | 'cow' | **buku** | 'book' | | |
| *Items for production of novel items*: | | | | | |
| **hippo** | 'hippopotamus' | ***tomat*** | 'tomato' | ***tomat*** | 'tomato' |
| katak | 'frog' | ***melon*** | 'watermelon' | ***melon*** | 'watermelon' |
| burung | 'bird' | **botol** | 'bottle' | **pizza** | 'pizza' |
| kupu | 'butterfly' | telur | 'egg' | telur | 'egg' |
| singa | 'lion' | mobil | 'car' | daging | 'bacon' |
| kalkun | 'turkey' | baju | 'shirt' | tali | 'rope' |
| lumba | 'dolphin' | tulang | 'bone' | | |
| rusa | 'deer' | surat | 'letter' | | |
| kura | 'turtle' | pohon | 'tree' | | |

*Note:* Bolded items are cognates; italicized items were presented using both the classifier for objects and the classifier for slices/pieces.

number, e.g. target *ta kar ayam* versus alternative *du kar ayam*, and in some cases a different classifier, e.g. target *du sum apel* versus alternative *du pak apel*.

A pilot study with real Indonesian words revealed unwanted differences in the length of the constructions. All of the nouns were bi-syllabic, as were Indonesian numerals (e.g. *ti.ga* '3') and classifiers (i.e. *e.kor, bu.ah, po.tong* for animals, objects and pieces), with the result that the reduplicated words were four syllables in length, while the numeral classifier constructions were six syllables. To remove this confound, we created artificial monosyllabic numerals and classifiers. This kept the length of the numeral classifier constructions at four syllables. To make sure that the numeral was not segmented as part of the frame, two different numerals were used: '2' and '3'. The artificial Indonesian forms for these were *du* and *ta* for '2' and '3'. The three classifiers for animals, objects, and slices or pieces were replaced by the nonce syllables *kar, pak*, and *sim* respectively. Thus, 'three apples' became *ta pak apel*, and 'two slices of bread' became *du sim roti*.[17]

## 3 Measures

We report participants' accuracy scores on trial1 and test, as well as the number of trials it took to learn to criterion for each linguistic variable. In order to demonstrate that learners retained what they had acquired, we report accuracy scores on re-test. We report production data from the NoPrompt Elicitation and the novel tasks. Because trial1

**Table 3.** Attempted production and accuracy per construction on the NoPrompt Elicitation task.

| Task | Frequency | Percentage of total | Accuracy (percent) |
|---|---|---|---|
| Bare | 200 | 53.1 | 170 (85.0) |
| Redup | 41 | 10.9 | 24 (58.5) |
| NumCl | 136 | 36.1 | 112 (82.4) |
| Total | 377 | | 306 (–) |

accuracy scores reflect the participants' performance on first exposure, data from participants who did not learn to criterion are also included. Only participants who learned to criterion provided data on the other measures.

### *4 Coding of production data*

Production data were coded as either 'attempted' or 'accurate'. Attempted responses were coded as Bare, Redup or NumCl regardless of whether the response was correct. A 'No Response' meant no attempt to produce a response. A second coding system separated the accurately produced responses from the inaccurate ones; for example, the substitution of one sound for another, say [ka.du] or [ku.ja] for Bare target [ku.da]. Since phonological accuracy is not the interest of this article, such errors were not counted as inaccurate attempts. Inaccurate attempts at reduplication involved mostly the irregular imitative reduplications. Participants produced the simpler full reduplication or used a wrong sound change rule. Errors of this sort show that learners were not merely imitating the input. Inaccurate attempts at NumCl arose from no or wrong use of the numeral, the classifier, or the noun.[18] Another error type involved using both Redup and NumCl (e.g. *ta pak buku-buku* to express 'three books').

## IV Results

### *1 Learning to criterion*

All 29 participants learnt to criterion on Bare, but only 22 participants did on Redup and NumCl. This provides one bit of evidence that learning the Bare construction was easy.

The mean number of trials to learn to criterion was 2.1 (*SD* 0.82) for Bare. Seven participants out of 29 learnt to criterion on the basis of one training trial; 13 participants did so within two training trials; all participants did so within four training trials. Thus, about two thirds of the participants learnt on the basis of two or four exposures (e.g. the training sentences + the verification questions). Learning the plurals was somewhat harder: the mean number of trials to reach criterion was on Redup was 3.27 (*SD* 1.64). Eight participants (out of 22) learnt to criterion on two trials; six participants within three trials. Thus, about two thirds of the participants needed six exposures to learn all items. The maximum required was seven trials (14 exposures). A paired $t$-test comparing the difference in the means of Bare and Redup was significant: $t(21) = -3.9$, $p < 0.001$, two-tailed. In learning NumCl, the mean was 4.09 trials (*SD* 2.79). Eight participants (out of

22) learnt to criterion within two trials; five participants within three trials. Thus, about half of the participants learnt on the basis of six exposures. The maximum required was 10 trials (20 exposures). A paired *t*-test comparing the difference between BARE and NUMCL was also significant: $t(21) = -3.83$, $p < 0.001$, two-tailed. Thus examination of the number of training trials needed to learn to criterion provides a second piece of evidence that learning the BARE construction was easier than learning the other two constructions: more participants learnt to criterion on fewer trials. It is also noteworthy that fewer trials were needed for even the slowest learner with BARE in comparison to REDUP or NUMCL. Finally, although it took on average more trials to reach criterion on NUMCL, the mean difference between REDUP and NUMCL was not significant ($t(20) = -0.84$, $p <$ 0.41, n.s.), meaning that NUMCL was neither harder nor easier to learn than the reduplicated plural.

## *2 Comparison of trial1, test and re-test*

Comparisons of scores on trial1, test and re-test permit an examination of performance over time and confirm that the linguistic variables had indeed been learnt. Comparison of performance by order of presentation of learning task (REDUP < NUMCL and NUMCL < REDUP) showed that the learning order had no effect.[19]

Twenty participants returned for the second session of the study. Examination of each individual's mean accuracy scores on trial1 and test of each linguistic variable revealed some outliers. The scores of these outliers were then adjusted to within two standard deviations of the means. Figure 1 shows adjusted mean scores, standard deviations, and frequencies of trial1, test, and re-test for each construction.

As Figure 1 shows, scores were high, even on trial1, and remained high after two weeks. One sample *t*-tests with the hypothesized mean set to 0.5 were calculated on all scores. Performance was above chance on all measures. For trial1 BARE: $t(28) =$ 33.69, $p < 0.001$, and test BARE: $t(28) = 46.25$, $p < 0.001$. For trial1 REDUP: $t(24) =$ 12.21, $p < 0.001$, test REDUP: $t(21) = 41.96$, $p < 0.001$, and re-test REDUP: $t(19) = 9.30$, $p < 0.001$. For trial1 NUMCL: $t(24) = 9.67$, $p < 0.001$, test NUMCL: $t(20) = 54.68$, $p <$ 0.001, and re-test NUMCL: $t(19) = 21.75$, $p < 0.001$. This provides evidence that learning had occurred.

To look at the development from one phase to the next, planned comparisons in the form of paired *t*-tests were computed on the means of trial1 and test for each construction. Differences in the means were significant: trial1 versus test for BARE, $t(28) = -2.72$ $p < 0.01$; trial versus test for REDUP, $t(20) = -5.63$, $p < 0.001$, trial1 versus test for NUMCL, $t(20) = -5.63$, $p < 0.001$, all two-tailed. This provides additional evidence that learning occurred during training. Paired *t*-tests of the differences in means on test and re-test produced significant differences: for REDUP $t(19) = 4.07$, $p < 0.001$; for NUMCL, $t(19) = 3.37$, $p < 0.003$, both two-tailed. These statistics show that there was forgetting in the two-week period between test and re-test.

A two-way Analysis of Variance (ANOVA) of linguistic variables (BARE, REDUP, NUMCL) and measures (trial1, test, re-test) was performed on participants' mean accuracy scores. Results were significant on both linguistic variables and measures, and the

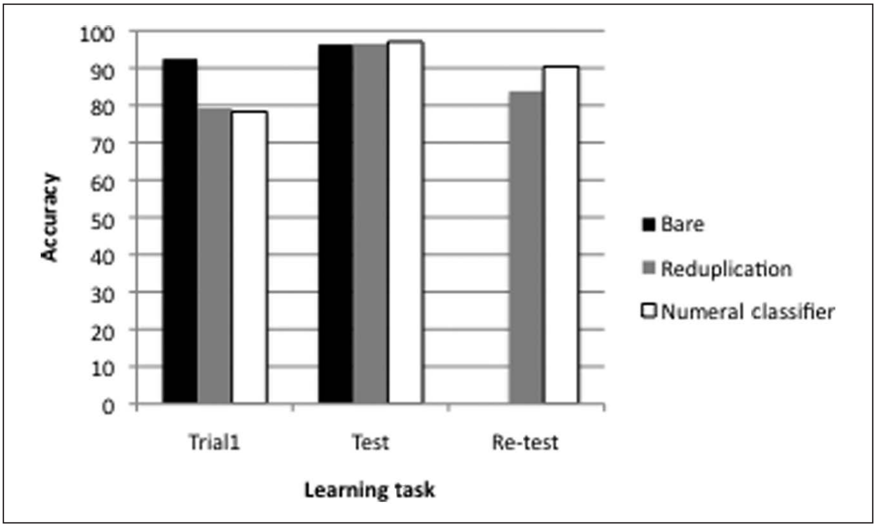



**Figure 1.** Mean accuracy scores for trial1, test, and re-test.

interaction between them was also significant. For the linguistic variables, $F_1(2,175) = 6.30$, $p < 0.02$; for the measures, $F_1(2,175) = 25.82$, $p < 0.001$; and for the interaction, $F_1(3,175) = 6.11$, $p < 0.001$. Using each test item's mean scores, a two-way ANOVA was also performed on the same variables. The results were again significant: For the linguistic variables, $F_2(2,152) = 6.78$, $p < 0.001$, and for the measures, $F_2(2,152) = 28.53$, $p < 0.001$. For the interaction, $F_2(3,152) = 6.28$, $p < 0.001$. This shows that performance on each linguistic variable was significantly different from the others at different developmental phases.

## 3 The production tasks

With data from 20 participants on 20 items, the maximum possible number of tokens produced would be 400. In fact, participants produced 377 responses. This amounts to a mean number of tokens per participant of 18.85 (out of 20). See Table 3.

Participants produced Bare over 50% of the time although they had been shown pictures of single (8 pictures) and multiple objects (12 pictures). These responses were consistent with the input and confirm the other evidence suggesting that the Bare forms were easy to learn. More importantly, they suggest that learners were not driven to interpret Bare forms as singulars (pace emergentism and FT/FA). Instead, this part of the data is consistent with the AIT, the Basic Variety Theory, the MTH and PT. However, and strikingly, over one third of the productions were NumCl, showing that learners were able to produce this construction. This is not consistent with the Basic Variety Theory, the MTH or PT. In comparison, the number of Redup plurals was small. One sample *t*-tests were calculated on the accuracy of the productions. Participants performed above

**Table 4.** Frequencies of production in obligatory and ambiguous contexts.

| Obligatory contexts (expected number of items) | Frequency of production (percentages in parentheses) | | | |
|---|---|---|---|---|
| | BARE | REDUP | NUMCL | No Response |
| BARE (8) | 134 (83.8) | 8 (5.0) | 9 (5.6) | 10 (6.3) |
| NUMCL (5) | 28 (28.0) | 5 (5.0) | 61 (61.0) | 6 (6.0) |
| Ambiguous contexts (7) | 38 (27.1) | 28 (20.0) | 66 (47.1) | 8 (5.7) |
| Total | 200 (50.0) | 41 (10.3) | 136 (34.0) | 23 (5.8) |

**Table 5.** Production frequencies on 24 novel stimuli (percentages in parentheses).

| | BARE | REDUP | NUMCL | Total |
|---|---|---|---|---|
| Attempted | 68 (14.5) | 99 (21.1) | 302 (64.4) | 469 (100.0) |
| Accurate | 62 (91.2) | 73 (73.4) | 240 (79.5) | 375 (79.9) |

chance on BARE and NUMCL, $t(199) = 13.83$, $p < 0.001$, and $t(135) = 9.86$, $p < 0.001$ respectively, showing that they could reproduce the forms relatively accurately. As for REDUP, production accuracy was at chance: $t(40) = 1.09$, $p = 0.28$, n.s. This was probably due to the difficulty in producing the imitative reduplications accurately. A chi-square goodness of fit test of the production accuracy of each attempted form was: $\chi^2(2) = 15.78$, $p < 0.001$. While the overall mean was quite high (81.2%), so was the standard deviation ($SD = 39.2\%$). One sample $t$-test revealed that five participants performed at chance. This was almost 25% of the sample size and suggests that initial representations were for these participants either inaccurate or unstable. This suggests that perhaps frequency of input may play a role in the accurate encoding of sound forms that are irregular in structure (compare Endress and Bonatti, 2007).

Performance on test items was also examined. One-sample $t$-tests on the mean accuracy of each stimulus revealed that three test items did not reach significance. These test items were the non-cognate words *bebek* ('duck'), *kuda* ('horse') and *kue* ('cake').

The NoPrompt Elicitation task included obligatory and ambiguous contexts. The obligatory contexts were ones where participants were expected to produce a given construction based on the input. Ambiguous contexts were those where any of the forms could be used appropriately. From these contexts, we can infer which forms learners accessed and/or computed more easily. Production data were categorized, based on the picture presented, to answer the question: How many times did the participants use BARE when they saw a single object? Recall that while the BARE construction was presented with pictures of both single and multiple objects, it was the only construction appropriate to represent single objects. We also asked: How many REDUP responses occurred when participants saw multiple items? Finally, how many NUMCL responses occurred when participants saw slices or pieces of objects? The NUMCL construction was the only one appropriate for rendering parts of objects.

Results are given in Table 4 which shows that participants produced all three constructions with pictures of single objects, but chose BARE almost 84% of the time. This use of BARE was (imperfectly) consistent with the input and clearly runs counter to a strict transfer story, namely if participants were biased to treat BARE as plurals (or as SUBSTANCEs) following English form–meaning mappings. If, however, BARE is easy to segment, store, recall and use (because its internal organization does not involve unification, pace PT), this result is predicted. A large proportion (28%) of BARE nouns was used where the NUMCL was expected. This is because three participants used BARE almost exclusively (20/20, 17/20, 19/20). Note that these responses were not wrong, but clearly show a response bias.

The ambiguous contexts were ones where only pictures of multiple objects were shown so that any of the three constructions could be used appropriately. Again, participants produced all three, suggesting sensitivity to the input. The large proportion of REDUP responses (20%) was only partly due to one participant who used it 19 times out of 20. No participant used NUMCL all the time, but four participants favoured this construction (12/19, 14/20, 13/20, and 13/19), which helps to explain why almost half the responses involved NUMCL. A chi-square goodness of fit test was calculated on the productions, and the result was significant: $\chi^2$ (6, $N$ = 400) = 147.59, $p < 0.001$, and, again, shows that participants had indeed acquired all three constructions.

The final task was the novel generalization task. Twenty participants were tested on 24 items; 470 tokens were recorded (maximum = 480). On average, participants made an attempt to produce 23.5 times out of 24 occasions. Table 5 summarizes the results and shows that participants produced all three constructions. Remarkably, almost two thirds were NUMCL, with another 21.1% being plural-marked forms. In addition, participants were accurate on all three. A chi-square goodness of fit test, calculated for the accuracy against the attempted production, $\chi^2$ (2) = 7.78, $p < 0.02$, showed that accuracy scores were well above chance. This shows that learners at the initial stage of learning are not constrained to produce only lexical items.

As with the NoPrompt Elicitation task, on the novel task stimuli were presented to test for particular responses. Four items were presented using pictures of single objects (favouring BARE). Five items were presented with pictures showing 4–6 objects (favouring REDUP). Six items were presented with pictures showing portions of objects (favouring NUMCL), 63/80 responses were provided in the expected BARE form (78.8%), 77/100 responses were in the expected REDUP form (77%), and 115/120 responses were in the expected form NUMCL form (95.8%). This provides strong evidence that the participants were sensitive to the input, had learnt each construction, and could generalize the patterns to novel stimuli. The remaining nine test items involved pictures of two or three objects and were compatible with all three response types. Three responses were in the BARE form (1.7%), 12 responses were in the REDUP form (6.7%), and 163 responses were in the NUMCL form (90.6%), showing that it was the preferred construction.

One sample $t$-tests were performed on the scores of the NoPrompt Elicitation task and the novel task. While five participants scored at chance on the NoPrompt Elicitation task, only two scored at chance on the novel task, showing that some learning continued to occur between the two tasks. Figure 2 shows the degree of improvement in production from NoPrompt Elicitation to novel.

Paired $t$-tests showed that these differences were not significant: BARE $t(17) = -2.00$, $p < 0.06$, n.s.; REDUP $t(9) = 0.99$, $p < 0.35$, n.s.; NUMCL $t(15) = -0.87$, $p < 0.39$, n.s. This is not surprising for the BARE and NUMCL constructions since the participants were performing extremely well already on the NoPrompt Elicitation task. It is also not surprising for the REDUP construction, given the phonological opacity of the imitative reduplicants,

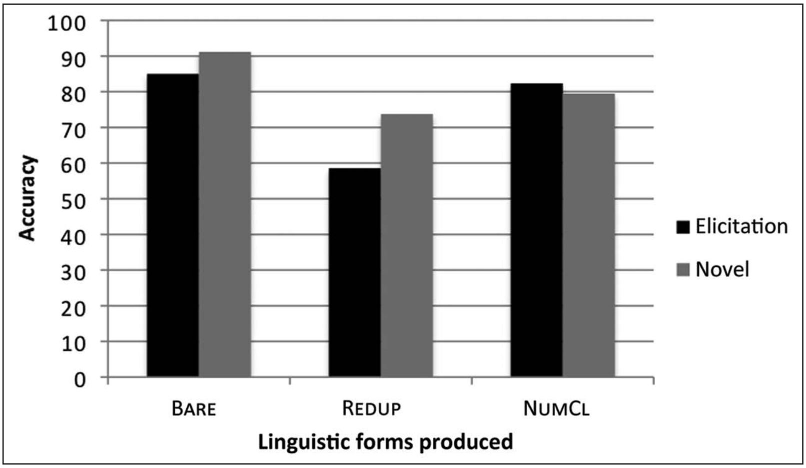



**Figure 2.** Accuracy of production on NoPrompt Elicitation and novel tasks.

which presumably require far more input for learners to cognize an abstract pattern in the variation. Finally a comparison of the accuracy of productions in Elicitation and novel for the ambiguous contexts showed a significant increase in the number of NUMCL forms produced in the novel task (from 60% to over 90%), $t(19) = -5.01$, $p < 0.001$.

## V Discussion and conclusions

The central challenge of second language acquisition research is to develop theories of language acquisition that explain patterns of acquisition, including constraints on those patterns. Different theories make distinct claims about the mechanisms involved, the role of input and the role of L1 transfer. We believe that the study of number – which cross-linguistically involves distinct patterns of mappings across phonological, morpho-syntactic and semantic representations – can shed light on the theoretical claims and promote useful discussion.

In an effort to understand what learners at the initial stage can learn given controlled input, we trained adult English speakers with no prior knowledge of Indonesian on three different constructions that can encode number in Indonesian, using pictures to promote picture–sound form referential mappings.[20] The constructions involved bare NPs, reduplicated plurals and a numeral + classifier + noun construction. Analysis of the results showed that performance on BARE was especially good. Even on trial1 of the training trials, performance was high, with low variance; participants took fewest number of trials to learn to criterion, and readily produced BARE forms in the NoPrompt Elicitation and novel tasks. Their production was remarkably accurate too. All the theories under consideration predicted good performance on the BARE construction, for different reasons. However, some of the theories – namely, emergentist theories and FT/FA – can be argued to make predictions about the interpretation of the BARE forms as well. A strict transfer hypothesis predicts that Anglophones should interpret Bare as plurals only since singular nouns in English require a determiner. While performance on the expected responses on the NoPrompt Elicitation task showed that participants were willing to use Bare to represent multiple objects, as required by the input, they clearly preferred it to represent single objects. This clearly demonstrates that the L1 is not 'filtering' interpretation, and suggests, as functionalists have long claimed, that learners may prefer one-to-one form–meaning mappings (Slobin, 1985) – in this case: Bare single object, Redup more than one object, NumCl enumeration of objects – despite what the input provides. In short, the strict transfer hypothesis of emergentism and FT/FA is not borne out for the Bare construction.

Contrary to the predictions of both the Basic Variety Theory and Organic Grammar, trial1 mean scores on Redup were above chance, suggesting that even after only two exposures, learners were mapping the plural forms to the correct meaning. However, participants exhibited greater variance and required more training trials to learn Redup. Nonetheless, participants performed at ceiling on test and were still significantly above chance on re-test. This was confirmed by the production data from the NoPrompt Elicitation task and novel task where participants were remarkably accurate. The 'remarkably' here is essential; the fact that learner productions were variable shows that they were not reproducing acoustic representations of the input ('imitation') but instead had internalized more abstract representations. In short, our results suggest that learners can readily represent and produce markers of Plural and there are no absolute constraints on early stages that prevent this.

Results show that learners also acquired NumCl. Performance was above chance on training trial1, and performance was at ceiling at test. While participants took longer to learn to criterion on this construction, they were very accurate in their production on the NoPrompt Elicitation task and favoured it in ambiguous contexts where all three constructions would have been consistent with the input. This was also the favoured construction with novel nouns. It is especially noteworthy that participants were not driven to produce input-inconsistent Numeral + Noun sequences as the Basic Variety and Organic Grammar theories predict they should have done. Moreover, the preference for the NumCl construction in ambiguous contexts where all three constructions were consistent with the input is entirely unexpected under these theories.

We believe that the predictions of the Basic Variety and Organic Grammar reflect particular properties of the languages studied, not universal constraints on early-stage grammars. We also believe that the investigation of functional categories in the early stages of SLA will not advance if researchers continue to believe that morpho-syntax can be studied without looking simultaneously at the acquisition of prosodic structure and the interactions between these two levels of representation. The study of such interactions is already shedding light on L1-based constraints on production in more proficient learners (Goad and White, 2006, 2008). There is every reason to believe that we will find L1-based constraints on prosody-morpho-syntax mappings at early stages.

We prefer a 'conservative' style of rhetorical presentation (Abelson, 1995) so we grant two important points. First, we have not shown that our learners have acquired the target functions of the Indonesian classifiers even if we have shown that they can creatively combine the classifiers with novel nouns. Second, we also do not claim to have shown that our learners have internalized the target structure of (9), although our approach permits us to formulate not only universal mapping correspondences such as 'Referent maps to NP' but also 'NUMERAL maps to NumPhrase' that would help to explain how learners could come to internalize such representations. It was not the purpose of our study to make such claims, and the acquisition of form–meaning mappings of classifiers merits much more research. Third, we acknowledge that it might be argued that our data show nothing more than an associative link between sound forms and meanings, with no clear evidence that learners have internalized a morpho-syntactic representation at all. This objection reflects a fundamental scepticism about the 'psychological reality' of morpho-syntactic representations that we feel we do not have to defend ourselves against, and it is certainly not a criticism that could be adopted by proponents of the Basic Variety Theory, Organic Grammar, or Processability Theory who believe, as we do, that learners can represent morpho-syntax from the earliest stages. We also note that the same complaint can be levelled at studies using spontaneous production data, the type of data that has served as the main evidence for these theories. Providing crucial evidence that learners have internalized autonomous morpho-syntactic representations independently of meaning and sound forms is very hard, perhaps impossible. Most research relies on logical argumentation instead. In our case, we point out that transfer hypotheses predict the ability to encode autonomous morpho-syntactic representations right from the start, as well as the ability to implement universal mapping strategies, such as the mappings from referents to NPs, and the universal X-bar theoretic constraint that all NPs must be headed by a noun. In addition, if we assume that learners begin L2 acquisition able to represent sounds and meanings but not morpho-syntax, we are faced with the impossible task of stating precisely when, how, and why, in L2 development, morpho-syntactic representation becomes possible (Fodor, 1980). Certainly, most of the theories that we have examined do not make the claim that learners cannot represent morpho-syntactic categories at the initial state; they differ in which categories they claim are representable.

Our data challenge claims that learners must follow the Processability Hierarchy since our participants had no difficulty in producing accurately NumCl, and performed worse on the reduplicated plurals. As for the AIT, it predicted that participants would learn all three constructions, and the data are consistent with this prediction. Nonetheless, they were less accurate in selecting the plural-marked Redup forms on the forced choice task and less accurate in producing them as well. We believe that this is due not to problems in representing the Redup construction per se, but rather to complications created by the use of imitative reduplications in the input and the use of minimal pairs in the forced-choice task. The use of minimal pairs in the forced choice task requires a shift of attention to acoustic detail that may make the task of selecting the target form much harder. The use of imitative reduplication as both targets and alternative choice forms probably increased the confusability of the target forms. Indeed, the fact that participants often chose a full reduplication alternative in the face of a target imitative reduplication suggests that they had internalized the full reduplication pattern. The question of the ease of learning reduplicated plurals vis-à-vis the numeral classifier construction merits another study with an improved design. Our experimenter error did have one saving grace; it underlines the (unsurprising) result that the phonological transparency of expressions plays a role in learners' willingness to (re)produce them. Finally, our investigation of Indonesian shows the value in exploring SLA with languages that have properties quite distinct from those of the western European languages that have served as the basis for the development of the various L2 theories discussed.

To conclude, we have shown that learners are not absolutely constrained by L1 properties in their analysis of L2 input involving three different ways to express number. Learners can readily learn a new interpretation for a familiar context (a bare NP); they can also learn several new forms for expressing a familiar grammatical category (Plural and the expression of numerals). In addition, they can do this on the basis of very little input. Learners are also not constrained to represent and produce only the sound–meaning mappings associated with lexical categories. They can also rapidly acquire sound–meaning mappings associated with functional categories, viewed here not as a distinction to be stated purely in morpho-syntactic terms, but rather as contrasts resulting from prosodic, morpho-syntactic and conceptual structures taken together (pace Muysken, 2008). Finally, our study adds to the literature showing that first exposure learners can segment and map L2 forms to pictures of referents but, significantly, shows that they can also analyse the senses of forms and start the process of learning L2 morpho-syntax.[21]

## Acknowledgements

We thank the reviewers for their questions and criticisms, which have greatly improved the final version of this article.

## Funding

This research was made| possible through generous funding to the first author by the Social Sciences and Humanities Research Council of Canada through its Canada Research Chairs program (950-202408).

## Notes

1. We ignore proper names in the rest of this article. Concepts are written in large capital letters throughout; names of Indonesian constructions are written in small capital letters.
2. We are reluctant to claim that FT/FA ‘predicts’ the early emergence of any novel construction because its authors are vague as to what language acquisition mechanisms are or how they interact with input.
3. So there is more involved than simply counting the frequency of forms in the input. Processors must be sensitive to such abstract notions as ‘register’ in order to encode the relative frequency of forms per text type. On the notion of register, see Gregory and Carroll (1978).
4. These representations are considerably simplified for expository purposes.
5. To keep the representations as simple as possible, we do not show the referential tier in (2) but discuss it here so that the reader will better grasp the assumptions we make about how learners deal with our word-learning tasks which presented them with pictures of objects.
6. Like Indonesian, Korean permits bare noun phrases to express non-specific reference. It may have a Plural-morpheme -*tul* but D MacDonald presents arguments that this morpheme serves a different semantic function in Korean. The fact that Koreans find English plural difficult suggests that they are not able to transfer a comparable morpheme to learning English.
7. Those familiar with formal semantic treatments of indefinite and definite determiners through the use of an existential or a uniqueness operator respectively (e.g. $\exists x$, BOOK(x); $\iota x$, BOOK(x)) will see here a radical departure from standard analyses. Jackendoff (1983: 13–6) assumes the ‘grammatical constraint’, namely that one will prefer a theory of semantics that captures generalizations about the grammar and lexis of a language (which what would

otherwise be arbitrary properties). In this case, the properties to explain are the widespread use of determiners to express definite and indefinite reference, and the general absence of constructions for this purpose like ‘there exists a thing and that thing is a book.’

8. In on-going research we have gathered robust evidence that Anglophones exposed to German for the first time have no difficulty perceiving, representing and producing syllables containing schwa or unstressed syllables when those syllables are analysable as part of an English Foot-structure, i.e. they occur as the second or third syllable following a strong syllable. Our participants could do this on the basis of minimal exposure to the words in question.
9. A reviewer points out that if English speakers are transferring L1 prosodic preferences they should distinguish lexical and functional words right from the start because Cutler and Carter (1990: 4–5) have shown that English speakers will treat a weak initial syllable as a functional category. This assumes, however, that the terms ‘lexical’ and ‘functional’ are specific to the morpho-syntax but we believe they should be understood in a ‘multi-dimensional’ way (Muysken, 2008: 5) as properties that result from mappings across the three types of representations. We say here in the text ‘when segmenting Prosodic Words’ that an English speaker analyses weak syllables as part of the content of the PW (as per (2c)). To analyse a weak syllable in ‘initial’ position as a Clitic means that there must be a PW as a sister node (as per (2a)). We predict that Clitics might be analysable initially only in utterance-initial position since the weak syllable cannot be analysed as anything else. Whether a stage 1 learner can analyse utterance-medial syllables as the weak left sister of a PW or are constrained (for whatever reason) to analyse only Prosodic Words remains to be seen. Certainly the literature on formulaic speech (Carroll, 2010 and references therein) suggests that learners are not initially treating weak syllables to the right of a PW as an Affix (as per (2b)). However, we grant that the right kind of experimental work remains to be done to put these predictions to the test.
10. *Me(m)* is a prefix that makes a verb of a root.
11. The original design called for the use of full reduplication only. The design was modified because the second author was led to believe that the contrast between regular and irregular forms necessarily correlates with on-line computation versus storage. It does not (see Baayen et al., 2002). By the time the error was discovered, it was too late to return to the original plan. As we will see, the use of imitative reduplication in the experiment obscures the main results.
12. As is typically the case in classifier languages, the sound forms of Indonesian classifiers map to distinct meanings: *orang* ‘person’, *ekor* ‘tail’ and *buah* ‘fruit’. We treat this as an indicator that there are distinct morphemes: one lexical, one functional. Thus *ekor* ‘tail’ is a noun, *ekor* indicating that its classified noun is an [ANIMAL], is a functional category.
13. A background questionnaire revealed that 23 participants were born in Canada, two in the UK, and one in the USA. The remaining three participants were born in the Colombia, Libya, and Philippines. English was the only childhood language for 25 participants, while four had English and another language as their childhood language.
14. Seven participants were monolinguals, 22 knew one to four other languages, eight were bilingual to some degree, 11 had knowledge of two other languages, two had knowledge of three other languages, and one had knowledge of four languages other than English. The reported languages were: Arabic, French, Greek, Italian, Spanish, and Twi. The questionnaire asked for any knowledge of other language(s), so it likely overestimates the real skills and abilities of the participants.
15. Although the current study was not interested in response latency, time constraints were applied to limit the length of the study. For Bare, participants had 2,500 ms to respond before the next item was presented; for Redup and Numcl, they had 5,000 ms to respond.
16. One of the reviewers believes that our results show ‘only’ that participants have relied on ‘chunking’. We have two comments. First, the term ‘chunking’ was introduced by Miller

(1956) to explain how a knowledgeable but non-fluent learner can develop fluent production (essentially by re-representing knowledge into larger constituents that increase the content of pronounceable sequences). In our study, learners were required first to segment forms and map them to meanings, and afterwards to creatively re-combine acquired units with novel ones. Their production errors show that they were doing just that. The question of how much learnt content is being re-encoded into larger units to increase fluency is not relevant here. Second, emergentists (e.g. Ellis, 1996, 2003; MacWhinney, 2005) have recently borrowed the term to describe segmentation and the parsing of sounds. This is what the reviewer appears to have in mind, meaning (apparently) that our learners have segmented acoustic sequences and are 'only' combining these to respond on the novel task. This is a very different notion from Miller's, and the production errors of our learners (e.g. combining REDUP with NUMCL which never occurred together in the input) suggest that they have more abstract representations in addition to acoustic ones. In any event, we believe that this dramatic shift in the meaning of 'chunking' as used by emergentists is a rhetorical strategy to disguise the fact that the segmentation and parsing of sound sequences involves learned prosodic units that are created through algorithmic processes. Certainly Miller has emphasized (1956: 92–93) that chunking in his sense requires 'a great deal of learning'. We believe that segmentation and the parsing of sounds do too, namely learning of the internal structure of (and constraints on) syllables, feet, prosodic words, and intonated phrases. And, yes, these units can be combined, but calling such combinatorial abilities 'chunking' tells us nothing about the cognitive processes involved. Nor are they trivial, so the 'only' is not warranted.

17. An acceptability judgement task was carried out with 26 native speakers of Indonesian, all from Jakarta, in order to ensure that they in fact responded to the three construction types as the grammatical literature said they should. Space restrictions prevent us from devoting discussion to this beyond saying that they did. Obviously, since we used nonce numerals and classifiers as our test stimuli we could not use native speakers as a control group.
18. Notice that omitting the numerals also shows that participants were not merely imitating the stimuli.
19. Two-sample *t*-tests were calculated on each task's measures. The differences were not significant. Group results were collapsed for further analysis.
20. One reviewer objects that our study lacks ecological validity. This objection can be levelled at virtually all research on speech perception and language processing. Nothing could be less ecologically valid than sitting with one's head strapped to an eye-tracker or being shoved into an fMRI scanner. Does this characteristic invalidate such research? We feel that the reviewer's objection simply misses certain fundamental points about scientific research. The first is: Is a particular method likely to lead to appropriate answers to one's research questions? In the case of those of us interested in understanding the nature of linguistic mental representations, highly ecologically valid research methods, such as ethnographic participant observation, have little to offer. Second, those of us who favour experimental methodologies in laboratory contexts do so in order to examine phenomena that cannot be easily studied in 'normal' learning situations. In particular, one goal of controlled input studies is to try to optimize learning to show what is possible. In this study, the density of input is undoubtedly considerably greater than it would be in a conversational context. If something is not acquired in our lab, it is unlikely to be acquired with less dense or more impoverished input. Third, those of us doing controlled input studies with first exposure learners are attempting to show that learners can represent linguistic distinctions rapidly, given input whose properties can be independently studied and described. This is not a result that can emerge from longitudinal corpus studies where there is no control of the input and where learners may have weeks, even months of exposure before they are first recorded. Finally, experimentalists understand

that they are engaged in a collective enterprise where evidence is gathered over time. We can challenge claims that certain kinds of phenomena must be absent from the initial state (as we believe we have done in this article) but we rarely 'prove' anything about language acquisition because, as the SLA literature shows, there is so little which is necessary and inevitable.

21. One reviewer says the following:

> Any theory that hypothesizes language-learning mechanisms has to recognize that these mechanisms are not the only means by which people can carry out language-related tasks. There is no reason why the participants in this study could not successfully treat these tasks as logical puzzles to be solved in ways that have little or nothing to do with language acquisition.

To which we reply: the claim is a logical non-sequitur so its 'truth' is not obvious at all. (1) Language-learning theories are responsible for postulating and describing language-learning mechanisms. They do not have to recognize anything at all about putative non-language-learning mental activities. (2) Not every theory discussed is even very clear about what the operations are of its language-learning mechanisms. Emergentist theories are; the AIT and PT are; the rest are not. (3) The theories that are explicit about their language-learning mechanisms differ radically, so it is glib to talk about 'any theory' as if they had to share common sets of assumptions. (4) The analysis of the processes involved in carrying out any linguistic task is difficult and involves a thorough understanding of speech perception (or reading processes where written language is involved), meaning mappings and meaning integration, and speech planning. We have some ideas about what it is our participants had to do; we are absolutely astounded that the reviewer claims to know what mental processes our participants must have been able to draw on to correctly respond to our tasks and to know what mental processes differentiate language acquisition from merely 'seeming' language acquisition activities. It is possible that the reviewer has in mind 'problem-solving', a term that shows up from time to time in the SLA literature without further discussion; for example, Bley-Vroman (1989) mentions it to motivate his Fundamental Difference Hypothesis. There is a literature on problem-solving that shows that it involves conscious reflection, reasoning in the form of inferences, and means-end analysis (Anderson, 1985; Carpenter and Just, 1999; Newell and Simon, 1972). Given the design of our study, in particular, the fact that participants were severely constrained to respond rapidly (within milliseconds), there is no reason to believe that they would have or could have adopted cognitive processes that exhibit these properties. Nor is there any reason to believe that such processes would have led to the accurate responses that our participants exhibited or even to their patterns of errors. We are well aware that some researchers feel that there is a motivated distinction between 'real' second language acquisition that takes place when people converse and 'not real' SLA that occurs in every other learning context. At the level of mental processing that we are concerned with, there is no empirical basis for such a distinction, as the neuro-linguistic literature is starting to show. Finally, second language learning, like most things humans do, must be fitted into models of rational behaviour. Yes, our participants were trying to carry out our tasks to the best of their ability: this is an example of rational behaviour and so are the efforts that learners make to attempt to understand what it is that a conversational partner is attempting to communicate.

## References

Abelson RP (1995) *Statistics as principled argument*. Hillsdale, NJ: Erlbaum.

Aikhenvald AY (2000) *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Allen K (1977) *Classifiers*. Oxford: Oxford University Press.

Anderson JA (1985) *Cognitive psychology and its implications*. 2nd edition. New York: WH Freeman.

Archibald J (1993) *Language learnability and L2 phonology: The acquisition of metrical parameters*. Dordrecht: Kluwer.

Baayen H, Schreuder R, de Jong N and Krott A (2002) Dutch inflection: The rules that prove the exception. In: Nooteboom S, Weerman F and Wijnen F (eds) *Storage and computation in the language faculty*. Dordrecht: Kluwer, 61–92.

Bailey N, Madden C and Krashen SD (1974) Is there a ‘natural sequence’ in adult second language learning? *Language Learning* 24: 235–44.

Barner D and Snedeker J (2005) Quantity judgements and individuation: Evidence that mass nouns count. *Cognition* 97: 41–66.

Bisang W (1999) Classifiers in east and southeast Asian languages counting and beyond. In: Gvozdanovič J (ed.) *Numeral types and changes worldwide*. Berlin: Mouton de Gruyter, 113–85.

Bley-Vroman R (1989) What is the logical problem of foreign language learning? In: Gass, SM and Schachter, J (eds), *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press, 41–68.

Bohn O-S and Flege JE (1992) The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition* 14: 131–58.

Carpenter PA and Just MA (1999) Computational modeling of high-level cognition versus hypothesis testing. In: Sternberg RJ (ed.) *The nature of cognition*. Cambridge, MA: MIT Press, 245–93.

Carroll SE (2001) *Input and evidence: The raw material of second language acquisition*. Amsterdam: John Benjamins.

Carroll SE (2004) Segmentation: Learning how to ‘hear’ words in the L2 speech stream. Special issue of *Transactions of the Philological Society: Empirical evidence and theories of representation in current research into second language acquisition* 102: 227–54.

Carroll SE (2010) Explaining how learners extract ‘formulae’ from L2 input. In: Véronique D (ed.) The processing of input in second language acquisition [Le traitement de l’input dans l’acquisition des langues étrangères’]. Special issue of *Language, Interaction, and Acquisition* 1 : 228–50.

Carroll SE (2012) Segmentation on first exposure to an L2: evidence for knowledge-driven, top-down processing. In: Braunmüller K, Gabriel C and Hänel-Faulhaber B (eds) *HSM13 multilingual individuals and multilingual societies*. Amsterdam: John Benjamins, 23–45.

Carroll SE (to appear, 2013) Processing ‘words’ in early stage SLA: A comparison of first exposure and early stage learners. In: Han Z-H and Rast R (eds) *Input processing in second language acquisition*. Cambridge: University of Cambridge Press, [page range needed].

Chung S (2000) Reference to kinds in Indonesian. *Natural Language Semantics* 8: 157–71.

Clark A (1993) *Associative engines: Connectionism, concepts, and representational change*. Cambridge, MA: Bradford Books/MIT Press.

Cutler A (1988) Prosody in situations of communication: Salience and segmentation. *Proceedings of the 12th International Congress of Phonetic Sciences* 1: 264–70.

Cutler A and Carter DM (1990) The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language* 2: 133–42.

Dalrymple M and Mofu S (2009) Plural semantics, reduplication and numeral modification in Indonesian. *Journal of Semantics* 29(2): 229–60.

Dulay HC and Burt MK (1973) Should we teach children syntax? *Language Learning* 23: 245–58.

Dupoux E, Pallier C, Sebastián-Gallés N and Mehler J (1997) A destressing ‘deafness’ in French? *Journal of Memory and Language* 36: 406–21.

Ellis NC (1996) Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition* 18(1): 91–126.

Ellis NC (1998) Emergentism, connectionism and language learning. *Language Learning* 48: 631–64.

Ellis NC (2003) Constructions, chunking and connectionism: The emergence of second language structure. In: Doughty CJ and Long MH (eds), *The handbook of second language acquisition*. Oxford: Blackwell, 63–103.

Endress AD and Bonatti LL (2007) Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition* 105: 247–99.

Ernst T (1991) A phrase structure theory of tertiaries. In: Rothstein S (ed.) *Syntax and semantics 25*. New York: Academic Press, 189–208.

Fodor JA (1980) Fixation of belief and concept acquisition. In: Piattelli-Palmarini M (ed.) *Language and learning: The debate between Jean Piaget and Noam Chomsky*. London: Routledge and Kegan Paul, 143–9.

Goad H and White L (2006) Ultimate attainment in interlanguage grammars: A prosodic account. *Second Language Research* 22: 243–68.

Goad H and White L (2008) Prosodic structure and the representation of L2 functional morphology: A nativist approach. *Lingua* 118: 577–94.

Goldschneider J and Dekeyser R (2001) Explaining the 'natural order of L2 acquisition' in English: A meta-analysis of multiple determinants. *Language Learning* 51: 1–50.

Gullberg M, Roberts L, Dimroth C, Veroude K and Indefrey P (2010) Adult language learning after minimal exposure to an unknown natural language. *Language Learning* 60: 5–24.

Gullberg M, Roberts L and Dimroth C (2012) What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics (IRAL)* 50: 239–76.

Gregory M and Carroll SE (1978) *Language and situation: Language varieties and their social contexts*. London: Routledge and Kegan Paul.

Hakuta K (1976) A case study of a Japanese child learning English as a second language. *Language Learning* 26: 321–51.

Hatch E (1983) Simplified input and second language acquisition. In: Andersen R (ed.) *Pidginization and creolization as language acquisition*. Rowley, MA: Newbury House, 64–86.

Jackendoff R (1983) *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff R (2002) *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jared D and Kroll JF (2001) Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language* 44: 2–31.

Klein W and Perdue C (1997) The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13: 301–47.

Larsen-Freeman D (1975) The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9: 409–30.

Li Y-hA and Shi Y (2003) NP as argument. In: Li Y-hA and Simpson A (eds) *Functional structure(s), form and interpretation: Perspectives from East Asian Languages*. London: Routledge, 3–27.

MacDonald D (2010) Second language acquisition of mass-count nouns by Koreans. Unpublished MA thesis, University of Calgary, Calgary, AB, Canada.

MacDonald R (1976) *Indonesian reference grammar*. Washington, DC: Georgetown University Press.

MacWhinney B (2000) Lexicalist connectionism. In: Broeder P and Murre J (eds) *Models of language acquisition: Inductive and deductive approaches*. Oxford, UK: Oxford University Press, 9–32.

MacWhinney B (2005) New directions in the competition model. In: Tomasello M and Slobin DI (eds) *Beyond nature-nurture: Essays in honor of Elizabeth Bates*. Mahway, NJ: Erlbaum, 81–110.

Marian V and Spivey M (2003) Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics* 24: 173–93.

Miller GA (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81–97.

Muysken P (2008) *Functional categories*. Cambridge: Cambridge University Press.

Newell A and Simon HA (1972) *Human problem-solving*. Englewood Cliffs, NJ: Prentice-Hall.

Pienemann M (1998) *Language processing and second language development*. Amsterdam: Benjamins.

Pienemann M (2007) Processability theory. In: VanPatten B and Williams J (eds) *Theories in second language acquisition*. Mahwah, NJ: Erlbaum, 137–54.

Rast R (2008) *Foreign language input: Initial processing*. Clevedon: Multilingual Matters.

Rast R (2010) The role of linguistic input in the first hours of adult language learning. *Language Learning* 60: 64–84.

Rast R and Dommergue JY (2003) Towards a characterisation of saliency on first exposure to a second language. In: Foster-Cohen S and Pekarek Doehler S (eds) *EUROSLA Yearbook 2003*. Amsterdam: Benjamins, 131–56.

Ritter E (1991) Two functional categories in noun phrases: Evidence from Modern Hebrew. In: Rothstein S (ed) *Syntax and semantics: Volume 25: Perspectives on phrase structure*. New York: Academic Press, 37–62.

Ritter E (1992) Cross-linguistic evidence for number phrase. *Canadian Journal of Linguistics* 37: 197–218.

Schwartz, BD and Sprouse, R (1994) Word order and nominative case in non-native language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In: Hoekstra, T & Schwartz, BD (eds), *Language acquisition in generative grammar*. Amsterdam: Benjamins, 317–68.

Schwartz BD and Sprouse R (1996) L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12: 40–72.

Shoemaker E and Rast R (2013, this issue) Extracting words from the speech stream at first exposure. *Second Language Research* 29.

Slobin DI (1985) Cross-linguistic evidence for the language-making capacity. In: Slobin DI (ed.) *The cross-linguistic study of language acquisition: Volume 2: Theoretical issues*. Hillsdale, NJ: Erlbaum, 1157–1256.

Vainikka A and Young-Scholten M (1994) Direct access to X′-theory: Evidence from Korean and Turkish adults learning German. In: Hoekstra T and Schwartz BD (eds) *Language acquisition studies in generative grammar*. Amsterdam: Benjamins, 265–316.

Vainikka A and Young-Scholten M (1996) Gradual development of L2 phrase structure. *Second Language Research* 12: 7–39.

Vainikka A and Young-Scholten M (2006) The roots of syntax and how they grow: Organic grammar, the basic variety and processability theory. In: Unsworth S, Parodi T, Sorace A and Young-Scholten M (eds) *Paths of development in L1 and L2 acquisition: In honor of Bonnie D. Schwartz*. Amsterdam: Benjamins, 77–106.

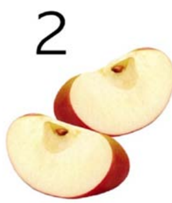
Weber A and Cutler A (2004) Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language* 50: 1–25.

Widjaja E (2010) Second language acquisition order of Indonesian reduplication and numeral classifiers. Unpublished MA thesis, University of Calgary, Calgary, AB, Canada.

Williams JN (2003) Inducing abstract linguistic representations: Human and connectionist learning of noun classes. In: van Hout R, Hulk A, Kuiken F and Towell R (eds) *The lexicon-syntax interface in second language acquisition*. Amsterdam: Benjamins, 151–74.

Yamada RA and Tohkura Y (1992) Perception of American English /r/ and /l/ by native speakers of Japanese. In: Tohkura Y, Vatikiotis-Bateson E and Sagisaka Y (eds) *Speech perception, production and linguistic structure*. Amsterdam: IOS Press, 155–74.

**Appendix 1.** Examples of visual stimuli.

| Task | Sample pictures | |
|---|---|---|
| Bare |  |  |
| Redup |  |  |
| NumCl |  |  |

## Reduplication stimuli

| Prompt | Forced choice (order randomized) |
|---|---|
| *Ini adalah ...* | *apakah ini ..., atau ...?* |
| 'these are …' | 'Are these … or …?' |

## Full reduplication targets

| | |
|---|---|
| ayam-ayam | ayam-ayam/ayim-ayam |
| bola-bola | bola-bola/boli-bola |
| coklat-coklat | coklat-coklat/ciklat-coklat |
| ikan-ikan | ikan-ikan/ikin-ikan |
| kuda-kuda | kuda-kuda/kudi-kuda |
| panda-panda | panda-panda/pandi-panda |
| pita-pita | pita-pita/piti-pita |
| zebra-zebra | zebra-zebra/zebri-zebra |

## Imitative reduplication targets

| | |
|---|---|
| apal-apel | apal-apel/apil-apel |
| baba-babi | baba-babi/babi-babi |
| balan-balon | balan-balon/bilin-balon |
| bebak-bebek | bebak-bebek/bebik-bebek |
| buka-buku | buka-buku/buku-buku |
| karta-kartu | kartu-kartu/kartu-kartu |
| keja-keju | keja-keju/keju-keju |
| koan-koin | koan-koin/koin-koin |
| kua-kue | kua-kue/kia-kue |
| leman-lemon | leman-lemon/limin-lemon |
| rota-roti | rota-roti/roti-roti |
| sapa-sapi | sapa-sapi/sapi-sapi |

## Numeral classifier stimuli

| Prompt | Forced choice (order randomized) |
|---|---|
| *Ini adalah …* | *apakah ini …, atau …?* |
| ‘these are …’ | ‘Are these … or …?’ |
| ta kar ayam | ta kar ayam/du kar ayam |
| ta kar ikan | ta kar ikan/du kar ikan |
| du kar panda | du kar panda/ta kar panda |
| du kar kuda | du kar kuda/ta kar kuda |
| ta kar babi | ta kar babi/da pak babi |
| du kar zebra | du kar zebra/du pak zebra |
| ta kar bebek | ta kar bebek/du pak bebek |
| du kar sapi | du kar sapi/du sim sapi |
| ta pak apel | ta pak apel/du pak apel |
| ta pak balon | ta pak balon/du pak balon |
| ta pak buku | ta pak buku/du pak buku |
| du pak kartu | du pak kartu/ta pak kartu |
| du pak bola | du pak bola/du kar bola |
| du pak koin | du pak koin/du kar koin |
| ta pak lemon | ta pak lemon/ta sim lemon |
| du pak coklat | du pak coklat/du sim coklat |
| du sim kue | du sim kue/ta sim kue |
| ta sim roti | ta sim roti/du sim roti |
| ta sim keju | ta sim keju/du sim keju |
| ta sim lemon | ta sim lemon/ta kar lemon |
| ta sim coklat | ta sim coklat/ta kar coklat |
| du sim pita | du sim pita/du pak pita |
| du sim apel | du sim apel/du pak apel |