LEARNING FUNCTIONAL CATEGORIES IN A SECOND LANGUAGE ON INITIAL EXPOSURE: CLASSIFIERS*

Susanne E. Carroll and Lindsay Hracs
University of Calgary

Abstract: We explore the interaction of linguistic and visual stimuli in the learning of nouns and classifiers in a novel language on first exposure. To interpret pictures, knowledgeable language users often rely on language that suggests what in the picture a speaker might be talking about. On first exposure to another language, this is not possible. It is often assumed that visual stimuli support inferences needed to learn the meanings of words. Within the Conceptual Semantics framework (Jackendoff, 1983, 2010, 2015), both noun phrases and nominal classifiers may express ontological categories such as THING, INDIVIDUAL, AMOUNT (of THINGs), (THING-)SHAPE, (THING-)SIZE, and (THING-)PROPERTY. Crucially, ontological categories may be independently accessed via visual stimuli to guide initial associations of conceptual representations and sound forms. We provide preliminary data showing that it is possible for adults to make such initial associations. Even with complex pictures, noun learning is comparatively easy. Classifier learning is much harder because it requires learners to extract "contrasts" across multiple stimuli.

1. Introduction

We study how adults, exposed to another language for the first time, come to learn words and the word classes they belong to.[1] Words express rich meanings in sentences, used in precise contexts. In such contexts, learners will learn the meanings of sentences and infer the contribution made by individual phrases and their heads, both those belonging to lexical categories (nouns and verbs) and those belonging to functional categories (like nominal classifiers). To interpret language requires that learners build morpho-syntactic structures (MSS). These are not isomorphic to semantic structures; indeed, they are processed in a distinct modular component of the language faculty (Jackendoff, 2010). However, units in

* We gratefully acknowledge generous funding from the Social Sciences and Humanities Research Council of Canada (grant number 435-2016-0087) which made this research possible. We thank Dr. Elias Abdollahnejad who took primary responsibility for creating and recording our Lab-Persian stimuli and for his native speaker expertise on real Persian. We thank the organizers of Colloque 50 ans de linguistique à l'UQAM : Regards croisés sur les enjeux de la linguistique for the opportunity to present a small part of our research, as well as participants for questions and commentary. Finally, we are grateful to anonymous reviewers for comments that led to important improvements to our original manuscript.

1 We capitalize ontological categories, e.g., THING; in syntactic representations word classes start with a capital letter, e.g. [Det [N]]NP; we use single quotes ‘’ for glosses and double quotes “” when citing and for terminology that is not well-defined or where there is no consensus on what the term denotes; we approximate lab-language sound forms as strings of International Phonetic Alphabet consonant and vowel symbols, e.g., /in_aest_mɛs.vɑk/ ‘this is scrub brush’ with the ‘_’ symbol displaying for the reader's benefit the word boundaries and the ‘.’ symbol displaying syllable boundaries.

the two systems "associate" in interesting ways which would be useful for a learner who must rely on the one to learn the other. Here we focus on correspondences between ontological categories and word classes.[2]

Learning novel words presents several challenges. We adopt the position, following Chomsky (1975: Ch. 2) and Pinker (1984), that the study of learning mechanisms is pointless in the absence of empirically-motivated theories of language. We also assume that learning mechanisms must be situated in theories of speech perception, prosodic and morpho-syntactic parsing, and interpretation (Carroll, 2001). Knowledgeable language users use representations of words, linguistic constraints, and rules of grammar to build the representations that make word recognition, sentence parsing, and interpretation possible. In the case of an unfamiliar language, the learner must create mental representations. While some aspects of language learning reflect universal constraints, whatever is not universal must be induced from the learning context. All languages have nouns that denote particular kinds of ontological categories (e.g., INDIVIDUALs, THINGs, SUBSTANCEs) but of course, languages differ in terms of the nouns of their lexica; nominal syntax varies in a number of ways cross-linguistically (see Section 2). For this reason, we are especially interested in the interaction and inter-dependency of linguistic and non-linguistic stimuli as forms of input causally related to learning outcomes.

Simple pictures of objects[3] (which make the "boundedness" of THINGs in space apparent) might encourage a bias to associate THING-concepts to sound forms. But speakers can talk about whatever takes their fancy, e.g, PROPERTY-concepts like the colour of the object, e.g., *This is yellow* when looking at a picture of a lemon or *This is wooden* when looking at a picture of a wooden bed, chair, or door or *This is round* when looking at a picture of a ball, a plate, or a bottle. Inferring properties of objects from pictures, rather than inferring the THING-ness of an object does not strike us as difficult, given the results of our previous research (Carroll, 2012, 2014; Carroll & Hracs, 2017; Carroll & Widjaja, 2013; Widjaja, 2010). However, associating such properties to the functional category Classifiers might not be easy. This is because the learner must look at different pictures of several objects, infer the **same property** is shared by all the objects in the set and, at the same time, cognize that the unique utterances that talk about the objects contain the same sound form, i.e., the Classifier (in Lab-Persian, /dʒɪn/ = Classifier$_{\text{WOODEN}}$, /tæn/ = Classifier$_{\text{ROUND}}$, and /ris/ = Classifier$_{\text{GRASPABLE}}$ (our SIZE classifier)).[4]

Complex pictures allow for a variety of form-meaning associations but learning word classes requires the learner to learn grammatically- and semantically-relevant "contrasts". Not all the concepts that are inferable from pictures can be expressed by nouns and nouns

[2] This assumption rests on the (de)compositionality of meaning where, e.g., noun phrases (NPs) correspond to arguments of predicates that, in turn, correspond to verbs and verb phrases (VPs) or adjectives and adjective phrases (AdjP). Not all sentences are decompositional; some are idiomatic. However, if all sentences were idiomatic, the functionalist assumption that "meanings" are the basis for learning syntax would necessarily be false. See Jackendoff (2015) for discussion.

[3] We have used simple pictures in several "first exposure" studies. See Carroll, 2012, 2014 for Lab-German, Carroll & Hracs, 2017 for Lab-Korean, and Carroll & Widjaja, 2013 and Widjaja, 2010 for Lab-Indonesian.

[4] Our choices of concepts came from linguistic descriptions of classifier languages (Aikhenvald, 2000; Downing, 1996; Greenberg, 1972; Grinevald, 2000; Matsumoto, 1993; Senft, 2000; and Unterbeck, 1994).

may be extraordinarily rich in comparison to the pictures. Classifiers express just the basic ontological category. The study of noun and classifier learning on first exposure is a good way to explore how adults learn individual words from pictures and abstract the relevant ontological category needed to cognize the generalization behind the use of the SHAPE-, PROPERTY-, and SIZE-classifiers.

2. Some grammar

2.1 English NPs

English NPs may be "bare" when headed by a proper noun (*Jane loves Bill*), a common noun head is Plural-marked and expresses a TYPE or class (*Gardeners plant flowers*) or denotes a SUBSTANCE or an AGGREGATE (*Jane ate pizza; Jane bought furniture*). When the NP is headed by a THING-denoting noun, English NPs obligatorily require functional categories like Determiners, Numerals, or Quantifiers (*Jane ate a pizza, two doughnuts, and drank several beers*). It is incorrect to conclude that the distribution of common nouns is distinct from proper nouns (as traditional grammars sometimes have claimed) because proper nouns can also be modified (*Every Fiona I know has red hair*) and at least some THING-denoting nouns also can have SUBSTANCE-interpretations (*Jane ate pizza and drank beer*). If #*Jane ate doughnut* is semantically anomalous, it is because a DOUGHNUT is a finished, baked artefact (a THING) and therefore does not lend itself to a SUBSTANCE-interpretation. Compare *Jane ate doughnut dough*.

To quantify SUBSTANCE-denoting Nouns, English relies on measurement phrases in lexical form: *a bit of sugar, a bowl of water, a cup of rice* (versus **a sugar, *a water, *a rice*). The function of such expressions is to "individuate" the SUBSTANCE-denoting Noun.[5] But English determiners {*a*, *an*, *the*} express far more than the boundedness of entities. They contrast in terms of dimensions like (in)definiteness and old/new information which are not relevant to the semantics of the target words.

2.2 Classifier languages

Cross-linguistic differences indicate what must be learned. Many languages like Indonesian, Japanese, and Lab-Persian lack determiners. In these languages, a novel referent expressed via a bare NP can denote both a singleton and multiple entities. Reference to THING-TOKENs may be expressed by demonstratives. See (1) and (2):

(1) a. Jane bought that book. (pointing to the referent)
    b. Jane bought a/*some book.

---

[5] These strings are syntactically ambiguous as shown by s(emantic)-selection. The verb crash requires a THING-denoting subject NP in A bowl of water crashed onto the floor. Listeners will infer that the head N is bowl, which means that of water is a complement of the N and the Determiner a determines the sequence, as in [a [bowl] [of water]]. In a bowl of water spilled onto the floor, the s-selection requirements of the verb spill tell listeners that the head N is *water* and *a bowl of* is a modifier with the Determiner determining *bowl*, as in [[a [bowl of]] water]. See Selkirk (1977). These examples show nicely why distributional analysis alone (without recourse to sentence meaning) cannot lead to the correct analysis of linear strings.

(2) a. An aenar xaerid-aem (Persian; see Gebhardt 2009)
DEM pomegranate bought-1^{st}SG
'I bought that pomegranate.' (pointing to the referent)

b. aenar xaerid-aem
pomegranate bought-1^{st}SG
'I bought a pomegranate.' or 'I bought pomegranates.'

To express measuring out and individuation, languages like Indonesian, Japanese and Lab-Persian rely on Classifiers. Classifiers can involve several ontological contrasts, e.g., HUMAN, ANIMAL, THING, and, as noted, they can express size and shape distinctions of THING-denoting nouns. Because pictures always "tokenize" the objects depicted, they may be quite precise in the size and shape information they provide as long as a visual context provides comparison objects. The contrasts classifiers encode are rather crude (e.g., small enough to be grasped in the hand vs. not, or overall shape of an object as long and thin). Classifiers can be category-specific (e.g., Thai has a classifier used with nouns denoting elephants). They can be homophonous to nouns, in which case the noun-form has a much richer meaning.[6] In many languages, classifiers are optional with respect to the head noun but may be obligatory in the presence of numerals and quantifiers. See (3):

(3) a. Se ta aenar xaerid-aem (Persian)
three CL pomegranate bought-1^{st}SG
'I bought three pomegranates.'

b. *Se aenar xaerid-aem
three pomegranate bought-1^{st}SG

c. *Ta aenar xaerid-aem
CL pomegranate bought-1^{st}SG

Sentences like those of (2) should tell the learner that NPs can occur without Demonstratives or Classifiers. The optionality of Lab-Persian classifiers was incorporated into our design by exposing participants first to bare NPs and only in Phase 3 to the Numeral+Classifier+Noun construction.

3. An overview of the learning problem

Linguistic theories model linguistic knowledge as formal descriptions. We think of our learners as having to create novel representations drawn from Conceptual Semantics and Simpler Syntax (Culicover & Jackendoff, 2005) within the tri-partite functional architecture of the human language faculty (Jackendoff, 2002, 2015). A learner who has learnt a word, category, or contrast can represent it in the appropriate representational system. A learner who has not cannot. Nonetheless, we gather data in the form of behavioural tasks and link tasks to representations by drawing on proposals by

[6] Gebhardt (2009) writes about classifiers as if they only occur in Numeral + Classifier constructions. In an overview, Senft (2000) provides examples of predicative classifiers which attach to verb roots or stems. We are concerned in this study with Numeral + Classifier distinctions only. How predicative classifiers are acquired remains to be investigated.

psychologists as to what learning mechanisms might be involved. Associative learning is a relatively unconstrained learning mechanism: "link-in-memory" X and Y. The view that words are associations of "sounds" and "meanings" is a commonplace in linguistics since Saussure (de Mauro, 1973). We adapt this view by hypothesizing that learners associate "sounds" as units of phonology to the units in a MSS. However, this association cannot inform the learner as to the grammatical identity and functions of the syntactic terminal node. For that, we hypothesize along with many developmental psycholinguists that learners are biased to associate a word class Noun to a THING- or to an INDIVIDUAL- or to a SUBSTANCE-concept (Clark, 1979; Pinker, 1984; Subrahmanyam, Landau & Gelman, 1999; inter alia). If pictures make the concepts available, then they can be exploited by such biases and learners will acquire their first nouns.

We hypothesize that noun-learning must precede classifier-learning so that individual sound forms have at least some ontological categories in the lexical entries of newly acquired Lab-Persian nouns. To learn the classifiers that co-occur with numerals and nouns, learners must cognize crucial differences between "lexical" word classes and "functional" word classes (Baker, 2003; Emonds, 2000; Muysken, 2008).

4. Methodology

We use a methodology that shows good reliability and versatility across several studies (Carroll, 2012, 2014; Carroll & Hracs, 2017; Carroll & Widjaja, 2013; Widjaja, 2010). In each study, our goal is to have a participant sample that is homogeneous with respect to the knowledge of the target language (they have none) and to exercise control over all linguistic and non-linguistic stimuli, as well as over the frequency and manner of exposure to the stimuli. This gives us more confidence when drawing conclusions about the causal role of input in SLA. We share such considerations with studies in which adults are trained on artificial languages (see Folia, Uddén, de Vries, Forkstam, & Petersson, 2010 and Hayakawa, Ning, & Marian, 2020 for reviews).[7]

The advantage of using "lab-languages" over artificial languages to create linguistic stimuli is that we can be reasonably sure that the target languages exhibit properties of natural languages, except for the explicit simplifications that we make.[8] Finally, although we report group data, we are mostly concerned with the emergence of knowledge in individuals. Mental grammars are the object of study; they do not exist in groups.

[7] We treat all studies in which adults are exposed to artificial languages as prima facie L2 studies even when the researchers believe their data shed light on first (L1) learning or assume that L1/L2 learning are identical (which, as the L2 literature shows, is clearly false).

[8] Lab-languages, including ours, are simplified in many ways: (i) they exploit very small lexica; (ii) they may reduce contrasts in functional categories or morpho-syntactic paradigms to one or two distinctions; (iii) they all appear to eliminate morpho-phonological irregularities of the English *go-went* type, as well as morpho-phonological variation that is manifested in form classes. While form classes are critical for learning the morphological structure of sound forms, they play no role in syntactic operations (Halle, 1992).

4.1 Participants

We tested 36 English-speakers,[9] all university students part of the Linguistics program participant pool. We excluded 10 participants from analysis for the following reasons: four participants failed to learn all target word/picture matches in Phase 1 (noun learning); three participants did not start learning English before the age of 3; three participants failed to provide responses for more than 3 stimuli items in a row. Statistical analyses were performed on data from 26 participants who provided response data on Phases 1, 2, and 3.

The mean age of the 26 participants was 19.9 years (range = 17–23). There were 23 participants who self-reported as female and 3 who self-reported as male. None reported knowledge of Persian, Kurdish or Arabic. Almost all had been exposed to at least one other language, either in the public school system or at university. All were exposed to English before the age of 3 and reported using English as a primary language of communication.

4.2 Materials and procedures

Before completing the study, participants were required to read and sign our ethics consent form and were asked to complete a biographical questionnaire and a computation span task to get a baseline measure of their working memory. We discuss these measures elsewhere.

Data were collected on computers in our laboratory using standardized materials and procedures. All learning tasks were run using PsychoPy2 (Peirce, Gray, Simpson, MacAskill, Höchenberger, Sogo, et al., 2019). Participants were tested individually. They were given instructions orally by a research assistant at the start of each task and also saw them in writing on the computer screen. All Lab-Persian linguistic stimuli were created and recorded by a native speaker of Persian (male voice). Participants listened to recorded material over studio-quality headphones. Visual stimuli consisted of still pictures selected from the children's Claymation series *Pingu* (Gutmann, the Pygos Group and Pingu Film Studio, 1986-1993) and were presented simultaneously on-screen with the appropriate linguistic stimulus presented auditorily.

Participants were exposed to target items in a sequence of training < test cycles. As noted, there were three phases: Phase 1 (Noun learning [13 trials]), Phase 2 (Numeral learning [9 trials]), and Phase 3 (Classifier learning [18 trials]). In both the training and test part of the cycle, participants were exposed to all stimuli one after the other with random presentation of utterances (matched to the relevant picture) throughout. In the training part of the cycle, visual stimuli remained on screen for a total of 5 seconds, and each stimulus presentation was separated by a fixation cross presented on screen for 1 second. In the test part of the cycle, visual stimuli remained on screen for the duration of the auditory stimulus plus 5

[9] We have learned that relying on self-reports of being a "native speaker" is a bad idea. We gather biographical information from participants, run everyone willing to participate, and discard the data of those who start learning English after the age of 3 (Meisel, 2011; see also Abrahamsson, 2012). A small number of students report having had exposure to two languages before the age of 3 years. We run their data and exclude any from final analyses if their response patterns suggest they are outliers.

seconds. Pilot testing revealed that 5 seconds post-stimulus was an appropriate time for participants to respond, taking into consideration the complexity of the visual stimuli.

The test involved a binary forced-choice procedure. Participants were instructed to respond by pressing a pre-identified key on a computer keyboard with QWERTY layout. The left key, Q (marked with the phrase "Option 1"), was to be used if the participant thought the correct choice was the first expression they heard in the linear sequence of forms, and the right key, P (marked with the phrase "Option 2"), if they thought the correct choice was the second expression. Two kinds of data were thus collected: accuracy scores and response latencies. There was a dependency in that if the participant failed to press a key within the allotted time, the response was automatically recorded as an error. Response latencies were analysed only on accurate responses.

At the end of each test cycle the proportion of correct responses was shown on the screen as feedback on the global score (not, we emphasize, on individual items). If the sum showed that the participant had made an error during the test part of the cycle, the training-test cycle began again and continued until the participant had learned all target forms (= "learning-to-criterion"). Participants were given a maximum of 10 trials to learn-to-criterion in each of the three learning tasks. Learning-to-criterion was a prerequisite to moving to the next phase. Participants who did not meet the criterion were excused from subsequent parts of the study but received full remuneration.

4.3 Design and stimuli

The study is a within-subjects design. The Lab-Persian lexicon is small: 13 Nouns, 3 Numerals and 5 Classifiers. All Lab-Persian words conform to English phonotactics.

We trained and tested our participants on Nouns denoting THINGs (individuated inanimate objects). To find suitable objects for our ontologically-defined classifiers ROUND, WOODEN, and GRASPABLE, we carefully examined our visual stimuli looking for multiple objects. Thus, we found pictures of balls, plates and bowls, pictures of wooden doors, wooden beds, and wooden chairs, and pictures of objects that would, in their conventional functions, be used by being grasped in a (human) hand (e.g., a scrub brush, a bottle, and a spoon). To capture the fact that natural languages sometimes require learners to learn a noun-specific classifier, we created a classifier /tup/ that was only used with the Lab-Persian word /mɑ.hi/ 'fish' (construed as a $[\text{FISH}]_{\text{SUBSTANCE}}$ or foodstuff, or as a $[\text{FISH}]_{\text{THING}}$ since it appeared as an intact object on a plate or on top of a soup bowl).

In Lab-Persian, classifiers only occur in Numeral+Classifier+Noun constructions. In each trial of Phase 3, target classifiers + nouns were preceded by /jɛ/ the numeral '1', /dɔ/ the numeral '2', and /sɛ/ the numeral '3', on which participants had been trained independently in Phase 2. Thus, each target noun appeared three times in the trial, each time with a different numeral. Each time the SHAPE, PROPERTY, and SUBSTANCE classifiers occurred (three times in a trial), they occurred with a different noun. Thus, the co-occurrence probabilities for the numerals and the classifiers changed across a trial. Since $\text{Classifier}_{\text{FISH}}$ occurred only with the FISH-denoting noun, /tup_mɑ.hi/ '$\text{Classifier}_{\text{FISH}}$ fish' appeared three times in a trial, preceded once by each of the numerals.

We also invented a default classifier /jær/. General classifiers occur in natural languages and L1 acquisition studies suggest that it is often the first one acquired, with the ontologically-specific classifiers emerging only later in development (O'Grady & Lee, 2006: 130). The default classifier occurred six times in a training/test trial, three times preceding a noun that occurred elsewhere with a SHAPE or PROPERTY/SUBSTANCE classifier, e.g., /sɛ_dʒɪn_sæn.dæl/ '3 Classifier$_{\text{WOODEN}}$ bed' and /dɔ_ jær_ sæn.dæl/ '2 Classifier$_{\text{default}}$ bed'. These nouns were thus repeated within the trial. Defaults also co-occurred with nouns that only co-occurred with the default, e.g., /dɔ_jær_ku.lɛ/ '2 Classifier$_{\text{default}}$ '(leather) bag/satchel'.

In the test cycle of Phase 3, the numbers and nouns were preserved in the conjoined sequences so that what differed was just the Classifier. However, on half of the items, the question presented a familiar, correct Numeral+Classifier+Noun sequence (e.g., /sɛ_dʒɪn_sæn.dæl/ '3 Classifier$_{\text{WOODEN}}$ chair') with an unfamiliar and incorrect sequence (e.g. /sɛ_tup_sæn.dæl/ '3 Classifier$_{\text{FISH}}$ chair').

To sum up, we constructed the classifier system first, based on our understanding of the classifier linguistic literature, and then chose nouns to match the appropriate ontological category based on what our visual stimuli would permit. We organized our trials to test hypotheses arising from the psychological literature concerned with learning mechanisms, on the one hand, and input patterns (frequency of occurrence; variability of co-occurrence of sound forms), on the other hand. Example stimuli from the training cycles can be seen in (4) and example stimuli from the test cycles can be seen in (5).

(4) a. /in_ æst_ **gɔl.dɑn** / (Phase 1: Noun learning)
this is **cactus**
'This is (a) **cactus**.'

b. /in_ æst_**sɛ**/ (Phase 2: Numeral learning)
this is **3**
'This is (the number) **3**.'

c. /in_ æst_**sɛ_jær_ gɔl.dɑn**/ (Phase 3: Classifier learning)
this is 3 CL$_{\text{default}}$ cactus
'These are **3 CL$_{\text{default}}$ cacti**.'

(5) a. /ɑ.jɑ in_ **gɛr.di_**æst_jɑ_æb.zɑr/ (Phase 1: Noun learning)
if this **ball** is or tool
'Is this (a) **ball** or are (these) tools?'

b. /ɑ.jɑ_in_ jɛ_ æst_jɑ_ **dɔ**/ (Phase 2: Numeral learning)
if this one is or **two**
'Is this one (thing) or **two** (things)?'

c. /ɑ.jɑ_in_ **dɔ tæn_kɑ.sɛ**_æst_jɑ_dɔ_tup_kɑ.sɛ/ (Phase 3: Classifier learning)
if this **2 CL bowl** is or 2 CL bowl
'Are these 2 CL$_{\text{ROUND}}$ bowls or 2 CL$_{\text{FISH}}$ bowls?'

The same visual stimuli that were used in the training part of the learning cycle were used in the test. Keep in mind that the visual stimuli we used were complex. There were plenty of objects a given sound form might have referred to. To focus attention on the numerical value of an expression, one, two, or three red arrows were overlaid on the scene to pick out the entities referred to. The participant's task was minimally to correctly match a segmented sound form to a picture. Relevant studies show that participants will learn to associate sound forms and "meanings" if given enough exposure whether or not linguistic stimuli are patterned (Carroll, 1999). But what is enough exposure? Prior research suggests that rote learning of arbitrary sound forms/meanings requires repeated exposure to stimuli. If learners readily learn to match linguistic stimuli to pictures right from the first items of Training Trial 1 it suggests that they are bringing useful prior knowledge to the task.

5. Results

There was considerable variability in accuracy on Trial 1 of the Noun learning task (Phase 1), but some participants were close to perfect ($n = 5$ or 19.23% of the sample). Over half the participants were matching the sound forms of noun utterances correctly by the end of Trial 3, replicating the results of our prior studies showing that mapping noun forms to pictures is easy. This is true despite the greater visual complexity of the pictures used here.

Slightly more than one-third of the participants (10/26) reached criterion on the Classifier learning task, and no one, among those that did, succeeded before Training Trial 4. The mean number of training trials for those 10 individuals was 6.50, approximately twice the number of trials needed for Noun learning. There was much greater variability with only one subject showing rapid accuracy (94.44% by Trial 3). Fifteen participants improved incrementally over several trials, reaching scores of 80% and above (whether or not they learned-to-criterion). In sharp contrast, when learning the Nouns, all participants ($n = 26$) were able to reach and maintain high levels of accuracy across the training trials, suggesting that they were not confusing traits in the pictures or sound forms.

Exact Binomial Tests were performed to ask if participants performed significantly above chance.[10] For the Noun learning task, results showed that 11/26 performed significantly above chance on the initial trial and 26/26 performed significantly above chance on the final trial. For the Classifier learning task, none of the participants performed significantly above chance on the initial trial and 19/26 participants (the majority) performed significantly above chance on the final trial. Findings indicate that although matching all linguistic stimuli to visual stimuli was hard, participants were nonetheless learning the linguistic stimuli, and repeated exposure made a difference.

Taken together, the data show that it was much harder for participants to match the linguistic stimuli to pictures on the Classifier learning phase than on the Noun learning phase. A summary of the above results is found in Table 1.

---

[10] Taking into account the number of trials we had ($n = 13$ for Noun learning and $n = 18$ for Classifier learning) along with chance being equal to 50% on a binary forced-choice task, participants must score 76.92% or higher to perform significantly above chance on the Noun learning task, $p = .046$, and 72.22% or higher to perform significantly above chance on the Classifier Learning Task, $p = .048$.

Table 1. Comparison of performance on Nouns and Classifiers

Measure	Noun Learning	Classifier Learning
accuracy scores for Trial 1 (%)	*M* = 67.75, *SD* = 19.70, range: 7.69–92.31	*M* = 52.13, *SD* = 8.76, range: 33.33–72.22
# participants to reach criterion (by trial)	Trial 2: 6, Trial 3: 14, Trial 10: 26	Trial 2: 0, Trial 3: 0, Trial 10: 10
# trials to reach criterion	*M* = 3.85 (*n* = 26), *SD* = 1.67, range: 2–9	*M* = 6.50 (*n* = 10), *SD* = 1.43, range: 4–9
# of participants who performed above chance	Initial trial: 11/26 (42.31%), Final trial: 26/26 (100%)	Initial trial: 0/26 (0%), Final trial: 19/26 (73.08%)

We also examined whether the category of the classifier was a predictor of accuracy scores and residual reaction times. Due to the correlated nature of our data, Generalized Estimating Equations were performed in R (R Core Team, 2021). Analyses showed that the category of the classifier is a significant predictor of accuracy score on the initial trial, $\chi^2(5) = 12.90$, $p = .024$, but not on the medial or final trial. Pairwise comparisons of estimated marginal means do not show any significant differences between scores by classifier on initial trials. Nonetheless, the highest accuracy scores are observed on items with the default and ROUND classifiers, while the lowest accuracy scores are observed on items that take the WOODEN classifier. This suggests there are two quite different factors involved: frequency of form vs. cognitive "accessibility" of the visual information. With respect to the frequency of the form, recall that the default classifier /jær/ occurred in 6/18 stimuli items, while the ROUND, WOODEN, GRASPABLE and unique classifiers each occurred on 3/18 stimuli items. Thus, better performance on the default classifier indicated sensitivity to the frequency of occurrence of the default across the utterances of the trial.

By the medial trial, the category of the classifier is a significant predictor of residual reaction times, $\chi^2(5) = 32.03$, $p < .001$, and final trial, $\chi^2(5) = 89.14$, $p < .001$ (but not the initial trial). Pairwise comparisons of estimated marginal means for the medial and final trials show that participants responded significantly faster on items that take the unique $\text{Classifier}_{\text{FISH}}$ than the other categories. This suggests that the high co-occurrence probability of /tup_mɑhi/ '$\text{Classifier}_{\text{FISH}}$ fish' facilitates recognition of the sound sequence.

In sum: (i) almost all participants were readily able to learn nouns and match them to pictures, despite the complexity of our visual stimuli in comparison to pictures used in our earlier studies. They segment and learn the nouns to criterion quickly in terms of number of exposures to training trails. Accuracy scores were high on the final trial and even those who failed to learn all items were wrong on a single item only. (ii) It was much harder to learn to pair the classifier construction to pictures. Scores were much lower across all training trials, and far fewer participants learned-to-criterion. Nonetheless, with repeated exposure to the stimuli, participants were performing significantly above chance, showing that learning to pair linguistic-to-visual stimuli (associative learning) is possible. (iii) On the initial trial, the category of the classifier is a predictor of response accuracy but not response latency, but on subsequent trials, classifier category is a predictor of processing speed but not response accuracy.

6. Discussion

It is widely assumed that visual stimuli in context will support language learning, especially at the initial stage. Our study adds to the evidence that this is possible when learners are learning nouns, replicating findings from Carroll (2012, 2014), Carroll & Hracs (2017), and Carroll & Widjaja (2013).

In contrast to other SLA studies of first exposure learning (Dimroth, 2018; Gullberg, Roberts, & Dimroth, 2012), our design ensures that the learners have learned the nouns and the numerals occurring in the Numeral+Classifier+Noun constructions. However, the fact the participants did not learn the classifiers-to-criterion is not consistent with the comparative ease in learning the classifiers of our Lab-Indonesian study or the study by Mueller, Hahne, Fujii & Friederici (2005) in which Germans learned a very simple classifier system of Japanese by playing a video game. Planned item analyses in future publications will, we hope, shed light on such differences.

In our research, we combine linguistic theoretical approaches to defining "learning problems" and psychological approaches to learning mechanisms (like associative learning, distributional learning, statistical learning) as investigated in studies in which adults are exposed to artificial languages. We are not alone in attempting to bridge the gap between approaches (Lidz & Gagliardi, 2015; Hracs, 2021; Lidz, Gleitman & Gleitman, 2003; Pinker, 1984; Yang, Ellman & Legate, 2015; inter alia) and there are numerous problems in investigating linguistic-theoretical matters in either perception and processing in knowledgeable language users or in learners (Phillips, Gaston, Huang, & Muller, 2012) but we note that the approach is better established in L1 research than in L2 research.

7. Concluding remarks

We have begun to explore how adults process linguistic and visual stimuli on first exposure to an unfamiliar language. More specifically, we have begun to think about how visual stimuli can be interpreted and how interpretations can help adult L2 learners "break into" the initial grammar of an unfamiliar language, learning lexical categories like nouns and functional categories like classifiers. Conceptual Semantics and the tri-partite functional architecture of language (Jackendoff, 2002, 2010, 2015) offer fertile soil in which to plant such seeds, offering a highly constrained vision of how "meanings" interact with grammar. Our paradigm has replicated findings found with other paradigms and shown itself to be reliable and flexible for studying the contexts in which linguistic input leads to learning. In our current project we have now extended this to visual stimuli.

References

Abrahamsson, Niclas. 2012. Age of onset and nativelike L2 ultimate attainment of morpho-syntactic and phonetic intuition. *Studies in Second Language Acquisition 34*: 187–214.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun Categorization devices*. Oxford: Oxford University Press.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.

Carroll, Susanne E. 1999. Adults' sensitivity to different sorts of input. *Language Learning 49*(1): 37–92.

Carroll, Susanne E. 2001. *Input and evidence: The raw material of second language acquisition*. Amsterdam: Benjamins.

Carroll, Susanne E. 2012. Segmentation on first exposure to an L2: Evidence for knowledge-driven, top-down processing. In *Multilingual individuals and multilingual societies*, ed. Kurt Braunmüller and Christoph Gabriel, 23–45. Amsterdam: Benjamins.

Carroll, Susanne E. 2014. Processing 'words' in early stage SLA: A comparison of first exposure and early stage learners. In *First exposure to a second language: Learners' initial input processing*, ed. Zhaohong Han and Rebekah Rast, 107–138. Cambridge: Cambridge University Press.

Carroll, Susanne E., and Lindsay Hracs. 2017. Learning Korean classifiers: A first-exposure study. Paper presented at EuroSLA 2017, University of Reading, Reading, UK.

Carroll, Susanne E., and Elizabeth (Nina) Widjaja. 2013. Learning exponents of number on first exposure to an L2. *Second Language Research 29*(2): 201–229.

Chomsky, Noam. 1975. *Reflections on language*. New York, N.Y: Penguin/Random House.

Clark, Eve V. 1979. *The ontogenesis of meaning*. Wiesbaden: Akademische Verlagsgesellschaft Athenaion.

Culicover, Peter, and Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.

de Mauro, Tullio. (Ed.). 1973. *Ferdinand de Saussure, Cours de linguistique générale, publié par Charles Bally et Albert Sechehaye, avec la collaboration de Albert Riedlinger*. Édition critique préparée par Tullio de Mauro. Paris: Payot.

Dimroth, Christine. 2018. Beyond statistical learning: Communication principles and language internal factors shape grammar in child and adult beginners learning Polish through controlled exposure. *Language Learning 68*(4): 863–905.

Downing, Pamela. 1996. *Numeral classifier systems. The case of Japanese*. Amsterdam: John Benjamins.

Emonds, Joseph E. 2000. *Lexicon and grammar: The English Syntacticon*. Berlin: Mouton de Gruyer.

Folia, Vasiliki, Julia Uddén, Meinou de Vries, Christian Forkstam, and Karl Magnus Petersson. 2010. Artificial language learning in adults and children. *Language Learning 60*(suppl.2): 188–220.

Gebhardt, Lewis. 2009. *Numeral classifiers and the structure of DP*. Doctoral dissertation, Northwestern University.

Greenberg, Joseph H. 1972. Numeral classifiers and substantive number: problems in the genesis of type. In *On language: Selected writings of Joseph H. Greenberg*, ed. Keith Denning & Suzanne Kemmer, 166–198. Stanford, CA: Stanford University Press.

Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. In *Systems of nominal classification*, ed. Gunter Senft, 50–92. Cambridge: Cambridge University Press.

Gullberg, Marianne, Leah Roberts, and Christine Dimroth. 2012. What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics (IRAL) 50*: 239–276.

Gutmann, Otmar, The Pygos Group, and Pingu Filmstudio. 1986–1993. *Pingu* series of animations.

Halle, Morris. 1992. The Latvian declension. In *Yearbook of Morphology 1991*, ed. Geert Booij and Jaap van Marle, 33–47. Heidelberg: Springer Verlag.

Hayakawa, Sayuri, Siqi Ning,and Viorica Marian. 2020. From Klingon to Colbertian: Using artificial languages to study word learning. *Bilingualism: Language and Cognition 23*(1): 7480.

Hracs, Lindsay. 2021. *Modelling exposure and input in language acquisition*. Doctoral dissertation, University of Calgary.

Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.

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

Jackendoff, Ray. 2010. The parallel architecture and its place in cognitive science. In *The Oxford handbook of linguistic analysis*, ed. Bernd Heine and Heiko Narrog, 593–614. Oxford: Oxford University Press.

Jackendoff, R. 2015. Genesis of a theory of language: From thematic roles (source) to the Parallel Architecture (goal) (Sort of an intellectual memoir). Ms., Tufts University.

Lidz, Jeffrey and Annie Gagliardi. 2015. How nature meets nurture: Universal Grammar and statistical learning. *Annual Review of Linguistics 1*: 333–353.

Lidz, Jeffrey, Gleitman, Henry, & Gleitman, Lila R. 2003. Kidz in the `hood: Syntactic bootstrapping and the mental lexicon. In *Weaving a lexicon*, ed. D. Geoffrey Hall, & Sandra R. Waxman. 603–636. Cambridge, MA: MIT Press.

Matsumoto, Yo. 1993. Japanese numeral classifiers: A study of semantic categories and lexical organization. *Linguistics 31*: 667–713.

Meisel, Jürgen M. 2011. *First and second language acquisition: Parallels and differences*. Cambridge: Cambridge University Press.
Mueller, Jutta L., Anja Hahne, Yugo Fujii, and Angela D. Friederici. 2005. Native and non-native speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience 17*(8): 1229–1244.
Muysken, Pieter. 2008. *Functional categories*. Cambridge: Cambridge University Press.
O'Grady, William and Sun-Young Lee. 2006. Classifiers and learnability: The role of recasts. In *Frontiers of Korean language acquisition*, ed. Jae Jung, 127–137. London: Saffron Books.
Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods 51*(1): 195–203.
Phillips, Colin, Phoebe Gaston, Nick Huang, and Hanna Muller. 2021. Theories all the way down: Remarks on 'theoretical' and 'experimental' linguistics. In *The Cambridge handbook of experimental syntax*, ed. Grant Goodall, 587–616. Cambridge: Cambridge University Press.
Pinker, Steven. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.
R Core Team. (2021). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
Selkirk, Elisabeth O. 1977. Some remarks on noun phrase structure. In *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 285–316. New York, NY: Academic Press.
Senft, Gunter. 2000. What do we really know about nominal classification systems. In *Systems of nominal classification*, ed. Gunter Senft, 11–49. Cambridge: Cambridge University Press.
Subrahmanyam, Kaveri, Barbara Landau, and Rochel Gelman. 1999. Shape, material, and syntax: Interacting forces in children's learning novel words for objects and substances. *Language and Cognitive Processes 14*(3): 249–281.
Unterbeck, Barbara. 1994. Korean classifiers. In *Theoretical issues in Korean linguistics*, ed. Young-Key Kim-Renaud, 367–385. Chicago, IL: University of Chicago Press.
Widjaja, Elizabeth (Nina). 2010. *Second language acquisition order of Indonesian reduplication and numeral classifiers*. Master's thesis, University of Calgary.
Yang, C.D., Ellman, A., & Legate, J.A. 2015. Input and its structural description. In A. Gallego & D. Ott (Eds.), *Fifty years later: Reflections on Chomsky's ASPECTS. MIT Working Papers in Linguistics*, pp. 285-292. Cambridge: MIT Press.