Verb Acquisition in English and Turkish: The Role of Processing

Maxwell Kon[1], Tilbe Göksun[2], Aslıhan Bagci[2], & Sudha Arunachalam[1]
[1] Boston University, [2] Koç University

Abstract

To determine the effects of processing load on verb acquisition within and across languages, we manipulated whether English- and Turkish-acquiring children aged 2.5 and 3 years heard novel verbs in contexts with a preceding easy-to-process unmodified subject determiner phrase, or a longer more difficult—but more informative—modified subject. While English learners succeeded only when no modifiers were present, Turkish learners performed at chance across the board, although the trend was for better performance in the no modifier condition as well. We propose that heavy subject DPs pose difficulty for verb learning, but that the point at which the DP becomes too heavy may differ depending on the language being acquired.

1. Introduction

Upon hearing an unfamiliar word, learners can draw from the linguistic context in which it appears to determine aspects of its meaning. This is particularly helpful for verb acquisition, as a verb's meaning bears strong relationships to the syntactic contexts in which it can appear, and the linguistic context can provide useful information about the verb's argument structure and selectional restrictions (e.g., Landau & Gleitman, 1985). The ability to use this information in the service of verb learning is known as syntactic bootstrapping, and it has been well established in children as young as 1.5 years of age (e.g., Arunachalam, Escovar, Hansen, & Waxman, 2013; Yuan, Fisher, & Snedeker, 2012).

But to make use of the syntactic bootstrapping mechanism, children must be able to parse the linguistic context effectively. This suggests a paradox: to acquire the meanings of verbs, children must parse the linguistic context in which they appear; but to become effective parsers of sentences, children must be able to generate structural representations of language online, and the structure of a sentence is often determined in large part by the verb.

Of course, partial linguistic representations may suffice for children to be able to benefit from syntactic bootstrapping. For example, if a child can successfully parse through the subject of a verb in an intransitive utterance, she is well positioned to posit a lexical representation for that new verb and make at least some inferences about its meaning, as the referential expression labeling

the subject may provide valuable information about what the verb means. A referential expression that uniquely picks out a specific referent will be most successful at helping children discover a new verb's meaning, as it will allow them to focus their attention on the specific event participant in the world being described. For example, if a group of children are playing on the basketball court, the utterance, "see the girl dribbling?" is not very helpful for the child who does not know the meaning of the verb "dribble". On the other hand, "see the girl with the pigtails dribbling?" is very helpful—assuming the child knows the noun "pigtails". But the more informative referential expression also requires more processing effort; it is longer, more syntactically complex, and evokes a richer discourse model, all of which may be expected to result in a heavier processing load.

We suggest, then, that for young learners there is a tradeoff between how easy the linguistic context is to process and how informative it is, particularly when it comes to using the linguistic context to acquire a new word embedded in it. Indeed, with respect to processing, there is evidence that speed of parsing the DP labeling the verb's subject determines children's abilities to acquire a novel verb downstream, as in the sentence, "The red car is on the *deebo*" (Fernald et al., 2008). And with respect to informativeness, there is evidence that children better acquire novel verbs when provided definite descriptions labeling the verb's arguments, as in "The boy is *pilking* a balloon" than pronouns, as in "He is *pilking* it" (Arunachalam & Waxman, 2011, 2015; see also Imai et al., 2005; Imai et al., 2008). Despite that these studies depicted only one actor and one object in the accompanying visual scene, such that the pronominal context was *in principle* just as informative, these studies found that the richer descriptions helped children identify the verb's referent. Thus, children benefitted from rich, and perhaps overly rich, semantic information. Perhaps redundant and overspecified linguistic expressions result in a more detailed or robust semantic representation that anchors children's efforts to construct a situation model (van Dijk & Kintsch, 1983). Redundancy is common in child-directed speech (e.g., Newport, Gleitman, & Gleitman, 1977), and since parents are skilled at accommodating their children's linguistic skill levels (e.g., Bellinger, 1980; Bornstein et al., 2007; Hoff-Ginsberg, 1994; Huttenlocher et al., 2007; Pan et al., 2005; Snow, 1972, 1977), we might imagine that they include such redundancy because they believe that it supports their children's comprehension.

The tradeoff between informativeness and processing cost is likely to differ across languages. For example, children acquiring languages like Korean and Japanese, that allow arguments to be dropped from the sentence, perform better with less information than more in similar paradigms (Arunachalam et al., 2013; Imai et al., 2005; Imai et al., 2008). One hypothesis to explain these findings is that children acquiring such languages have less experience with—and thus more difficulty parsing—contexts with richer information. As of yet there is minimal evidence exploring this hypothesis, and the evidence that does exist comes only from languages of East Asia; we aim to provide new evidence on this issue from Turkish, another well-studied language that allows argument

drop. In Turkish, both subjects and direct objects can be dropped and word order is not needed to indicate thematic relations. Rather, case marking (e.g., accusative case marking) may be a reliable cue for learners.

A clear research agenda presents itself: To understand how children use linguistic context in verb acquisition, we must investigate the tradeoff between informativeness and processing cost both within and across languages, investigating different kinds of learning tasks (e.g., acquiring novel verbs with one, two, or three arguments; acquiring verbs given a crowded visual scene with many possible referents vs. few possible referents). We take a first step in the current study by asking how English- and Turkish-acquiring children acquire novel verbs when given varying levels of informativeness and processing cost in the accompanying utterances. Specifically, we presented children with novel verbs in intransitive frames that either had a simple definite description labeling the subject (e.g., "The girl is *fezzing*") or a more informative but longer and more complex description (e.g., "The nice tall girl is *fezzing*").

We also contrasted the visual scenes children were presented with when hearing the utterance in order to manipulate whether the longer description was informative or not given the context: the visual scenes either presented one possible referent for the noun in the DP or two possible referents—in the latter case a modifier was crucial for distinguishing between them. This resulted in three conditions. In the 0 Adjective/Different condition, children heard an unmodified subject DP and saw scenes with two different kinds of event participants (e.g., a girl, a boy). In the 2 Adjective/Different condition, children heard a modified subject DP and also saw scenes with two different kinds of event participants. In the 2 Adjective/Same condition, children heard a modified subject DP but saw scenes with two event participants belonging to the same category, with some salient property distinguishing them (e.g., a tall girl, a short girl). (The fourth logical possibility, in which two adjectives are presented but the scenes depict the same kind of event participant, would be unresolvable because some kind of modifying information is required for disambiguation; the noun alone does not suffice.)

Overall, the task we designed was quite difficult in that it presented the novel verb only once in an informative linguistic context. Success required several feats: 1) comprehending the DP in subject position, 2) quickly identifying which of the two scenes depicts the referent of that DP, 3) determining the likely referent of the verb based on the event occurring in that scene, and 4) choosing that referent again given a new exemplar at test. Of these, processing considerations should be most evident in step (1) in the 2-adjective conditions, because the 2-adjective sequence imposes a heavier processing load than the DP without adjectives. For Turkish learners, we anticipated one of two outcomes. It might be that even the unmodified DP is difficult for young Turkish children to process; given the availability of argument drop in Turkish and parents' tendency to talk about referents that are accessible and hence able to be dropped, Turkish learners are likely to hear fewer utterances with overt subject DPs and may thus have difficulty processing

them. Indeed, studies of verb learning in children acquiring East Asian languages that allow argument drop do suggest that even pronouns and unmodified definite descriptions may be difficult for children to process, at least for transitive verbs (Arunachalam et al. 2013; Imai et al., 2005, 2008). Alternatively, it could be that because of the availability of argument drop in appropriate discourse situations, children acquiring Turkish, more so than children acquiring English, are learning to be eminently sensitive to discourse factors governing whether modifiers are or are not appropriate in a given situation. Thus they might be worse off than English learners in the overinformative 2 Adjective/Different condition and better off than English learners in the 2 Adjective/Same condition.

2. Method

Participants. Native English-speaking children recruited from Boston, MA and native Turkish-speaking children recruited from Istanbul, Turkey participated (each group N = 48). The mean age of both groups was 3;3 (English learners' age range: 2;6 to 3;11; Turkish learners' age range: 2;5 to 4;0). Two additional English learners and six additional Turkish learners were tested but excluded from the final sample because they pointed on two or fewer trials.

Materials and Design. We introduced a forced-choice video pointing task in which children viewed four trials, each consisting of a familiarization phase and a test phase, and presenting a different novel verb. The design was between-subject: participants were assigned to the 0-adjective Different condition, the 2-adjective Different condition, or the 2-adjective Same condition.

On each trial, the familiarization phase depicted either two actors or two toys, each performing a different action (e.g., clapping, waving); the two video clips played simultaneously, side-by-side. In the Different conditions, these event participants were members of different basic-level categories (e.g., a girl, a boy). In the Same condition, these event participants were from the same basic-level category (e.g., two girls). In all conditions, the two event participants differed along some salient dimension (e.g., tall girl, short boy). The audio stimuli incorporated a novel verb in either a Det-N or Det-Adj-Adj-N phrase, depending on condition (e.g., "The (nice tall) girl is fezzing" in English; "(Uzun güzel) kız feziyor" in Turkish).

At test, a different event participant performed each of the two actions. The audio stimuli directed children to point to the scene depicting the referent of the novel verb (e.g., "Where is fezzing?" in English; "Nerede feziyor?" in Turkish). See Figure 1.
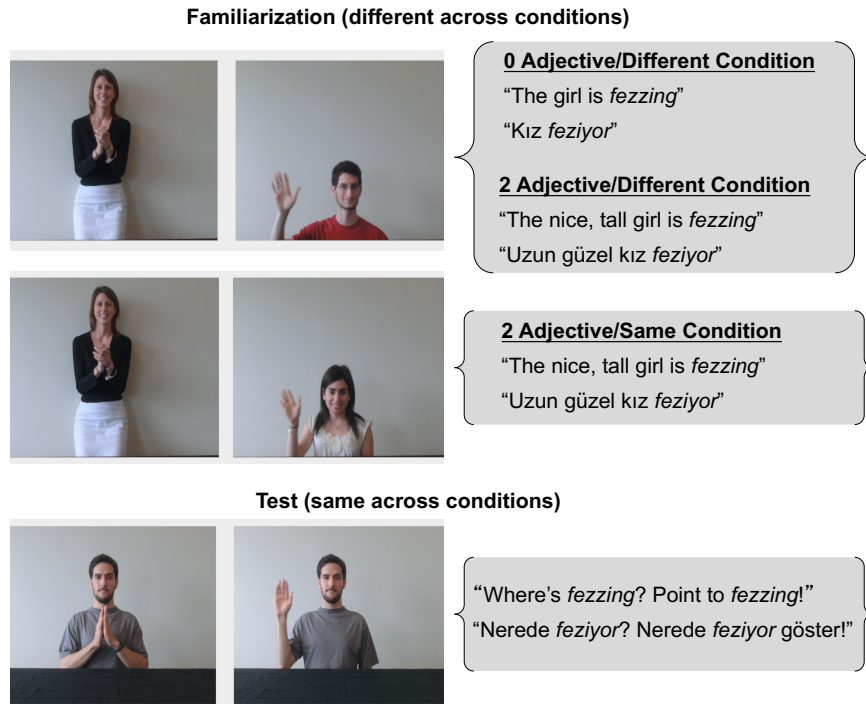
Familiarization (different across conditions)

0 Adjective/Different Condition

"The girl is *fezzing*"

"Kız *feziyor*"

2 Adjective/Different Condition

"The nice, tall girl is *fezzing*"

"Uzun güzel kız *feziyor*"

2 Adjective/Same Condition

"The nice, tall girl is *fezzing*"

"Uzun güzel kız *feziyor*"

Test (same across conditions)

"Where's *fezzing*? Point to *fezzing*!"

"Nerede *feziyor*? Nerede *feziyor* göster!"

Figure 1. Sample trial (one of four)

Procedure. Videos were presented to children on a computer monitor. We first presented the children with two warm-up trials, which were designed to encourage pointing to one side of the screen. Two video clips of familiar characters (Sesame Street characters for the English sample, farm animals for the Turkish sample) were presented side-by-side on the screen, and the experimenter asked the participant to point to a scene described with a linguistic prompt—once to a particular character (e.g., pig vs. cow), and once to a particular action (e.g., dancing vs. eating). On warm-up trials only, the experimenter demonstrated the correct response if the child pointed incorrectly. The warm-up was followed by the four experimental trials. At test, they were asked to indicate their choice of scenes by pointing, and no feedback was provided.

Coding and Analysis. For both English- and Turkish-acquiring children, pointing responses were recorded by two trained coders, naïve to study hypotheses. One coder was seated next to the child and elicited pointing; the other sat behind a curtain and recorded pointing responses via real-time video feed. Disagreements between coders were resolved by discussion.

Pointing was coded as a binary response; the child's first point was coded as a 1 if directed to the correct scene and a 0 if directed to the other scene. Trials on

which children did not point at all were excluded; this affected three trials from three different English learners, and no trials from Turkish learners. Because there are two scenes, chance performance is a 0.5 proportion of points to the correct scene.

Predictions. We entertained three outcomes: (1) If children benefit from more semantic information than is strictly speaking necessary to identify the referent, then we would predict success in the 2-adjective conditions whether or not children succeed in the 0-adjective condition. (2) If, however, children are pragmatically adept enough to recognize the overinformativeness in the 2-adjective Different condition, and be negatively affected by it, they may succeed in the 2-adjective Same condition but not the 2-adjective Different. (3) But if the 2-adjective DPs simply pose too great of a processing burden, then we predict the best performance in the 0-adjective condition.

The properties of the language being acquired are relevant for determining the likelihood of these possibilities. If Turkish learners have difficulty processing even unmodified definite descriptions, then the first potential outcome is particularly unlikely. If, on the other hand, they can process unmodified definite descriptions but have particular difficulty with overinformative referential expressions, then the second possibility is most likely.

3. Results

The results are presented in Figure 2. English learners performed well in the 0-adjective condition, selecting the target scene 67% of the time; they did so only 56% of the time in the 2-adjective Different condition and 51% of the time in the 2-adjective Same condition. Turkish learners showed a similar trend, though performance was poorer in the 0-adjective condition than in the English sample; they selected the target scene 57% of the time in the 0-adjective condition, 51% of the time in the 2-adjective Different condition, and 49% of the time in the 2-adjective Same condition.

We submitted the data to mixed-effects models (binomial family) using the glmer() function in the lmer package in R (v.3.2.1) (Bates & Bolker, 2012; R Development Core Team, 2012). To test significance, we used the z-test and p-values output by glmer(). Because our question of interest was whether children successfully chose the target scene at above chance levels, we tested each condition separately in a model including subject (as a random intercept) as well as child's age (centered around the mean) and child's sex and assessed whether the intercept parameter significantly differed from 0 (chance). We found that for English learners, only in the 0-adjective Different condition was performance significantly above chance (intercept $p = 0.0037$). In the other two conditions, performance did not differ from chance. For Turkish learners, no differences from chance were found. See Table 1 for model parameter estimates.

Figure 2. Proportion of pointing responses directed to the target scene, grouped by linguistic condition and language group
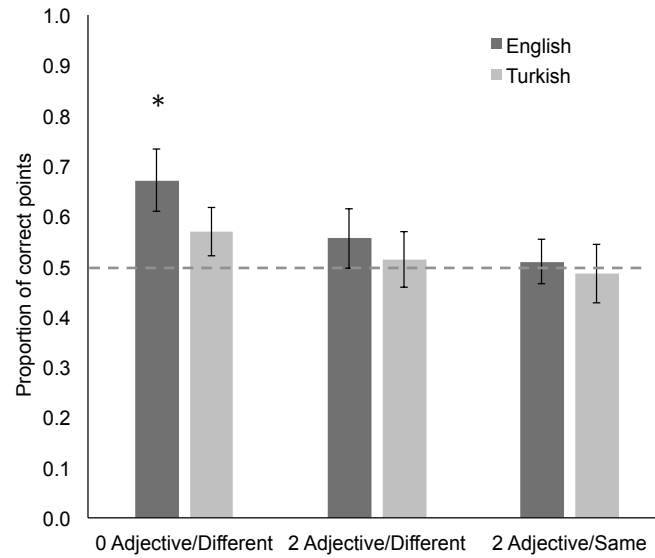

Table 1. Fixed effects from all models.

Language	Condition	Effect	Estimate	SE	z-value	*p*-value
English	0-adjective Different	Intercept	0.88*	0.30	2.90	0.0037
		Age	0.057	0.053	1.077	0.28
		Sex	-0.34	0.30	-1.12	0.26
	2-adjective Different	Intercept	0.21	0.24	0.90	0.37
		Age	0.0097	0.066	0.015	0.99
		Sex	-0.088	0.24	-0.36	0.72
	2-adjective Same	Intercept	0.034	0.25	0.13	0.89
		Age	-0.077	0.061	-1.26	0.21
		Sex	-0.038	0.26	-0.14	0.89
Turkish	0-adjective Different	Intercept	0.31	0.24	1.28	0.20
		Age	0.037	0.064	0.58	0.56
		Sex	-0.22	0.24	-0.92	0.36
	2-adjective Different	Intercept	0.056	0.24	0.23	0.81
		Age	-0.0034	0.059	-0.059	0.95
		Sex	-0.061	0.25	-0.24	0.81
	2-adjective Same	Intercept	0.0070	0.24	0.029	0.98
		Age	0.038	0.054	0.70	0.49
		Sex	0.29	0.25	1.17	0.24

* indicates p < 0.01

These results best support the third of the possibilities we laid out, that the 2-adjective DP presents such a great processing burden that verb acquisition is hindered, even though more information is included in these longer DPs. With respect to our cross-linguistic hypotheses, we did not find support for or against our hypotheses about cross-linguistic differences between English and Turkish learners. We did find that the Turkish learners performed worse in the 0-adjective condition, which could be taken as evidence in support of our hypothesis that even an unmodified DP is too much to process; we would predict that they would perform better with pronouns, for example, which could convey crucial gender information but without the processing load imposed by a definite description. We suggest this interpretation only tentatively, however, because the fact that we have only null results from the Turkish learners means that we cannot rule out the possibility that our stimuli or task were not well suited to this population.

4. General Discussion

Because of the difficulty of acquiring verb meanings from observation of the world alone (e.g., Gleitman et al., 2005), a linguistic context that helps learners identify the event being referred to may be extremely important. But importantly, when considering what makes a helpful linguistic context, we cannot simply look at the information adults find helpful; we must consider that children's intake of the linguistic information they hear is both limited by the capacities of their language processing system and may reflect a greater need than adults for rich semantic information in order to establish meanings for unfamiliar words. In this study, we asked whether the acquisition of novel verbs in intransitive frames in English and Turkish would be affected by the addition of adjectival modifiers to the subject DP, and if so, whether the effect would be positive—because additional semantic information is encoded, or negative—because the processing burden is increased. Our results, that English learners succeeded only when no modifiers were present and Turkish learners did not succeed in any condition (although they also performed best when no modifiers were present) suggest that the processing burden of two adjectives outweighed any possible benefits of added semantic information those adjectives provided.

Our overarching goal in this research program is to identify the features of a linguistic context that make it more or less successful in supporting verb acquisition, as well as how these features interact with each other. We aim to identify optimal "windows" of the type of linguistic context that supports learning in various contexts and in different languages.

Although decades of research have highlighted children's abilities to learn new word meanings from their linguistic contexts by combining their knowledge of syntactic structure, lexical semantics, and the extralinguistic context, the present study documents a limitation in these abilities. It is not simply the case that the more information children are given, the more successful they are at

acquiring new word meanings. Rather, the information must be packaged such that it can be easily processed by the child's developing parser. It may seem obvious to say that children can only learn from information that they can process. But we think this is an important point. First, this processing requirement has not been adequately acknowledged in the prior literature. (This paper's last author ended a paper in the proceedings of an earlier BUCLD meeting with a sentence that we strongly qualify, if not directly contradict, in the present paper: "The more [semantic] information [children] have, the more successful they are at deducing the meaning of the words they are learning" (Arunachalam & Syrett, 2014).) Second, identifying optimal linguistic contexts for learning will be invaluable for guiding research aimed at helping parents, clinicians, and educators support vocabulary learning in children with or at risk for language delays and disorders.

References

Arunachalam, Sudha, Escovar, Emily, Hansen, Melissa A., & Waxman, Sandra R. (2013). Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and cognitive processes*, *28*(4), 417-425.

Arunachalam, Sudha, Leddon, Erin M., Song, Hyun-joo, Lee, Yoonha, & Waxman, Sandra R. (2013). Doing more with less: Verb learning in Korean-acquiring 24-month-olds. *Language acquisition*, *20*(4), 292-304.

Arunachalam, Sudha, & Syrett, Kristen. (2014). Specifying event reference in verb learning. In: William Orman & Matthew J. Valleau (Eds.), *Proceedings of the 38th Annual Boston University Conference on Language Development* (pp. 39-50). Somerville, MA: Cascadilla Press.

Arunachalam, Sudha, & Waxman, Sandra R. (2010). Specifying the role of linguistic information in verb learning. In: Katie Franich, Kate M. Iserman, & Lauren L. Keil (Eds.), *Proceedings of the 34th Annual Boston University Conference on Language Development* (pp. 11-21). Somerville, MA: Cascadilla Press.

Arunachalam, Sudha, & Waxman, Sandra R. (2011). Grammatical form and semantic context in verb learning. *Language learning and development*, *7*(3), 169-184.

Bates, Douglas M., & Bolker, Ben. (2012). lme4: linear mixed-effects models using S4 classes. R package v. 0.999999-0. See http://cran.r-project.org/web/packages/lme4/index.html.

Bellinger, David. (1980). Consistency in the pattern of change in mothers' speech: Some discriminant analyses. *Journal of child language*, *7*(03), 469-487.

Bornstein, Marc H., Hendricks, Charlene, Haynes, O. Maurice, & Painter, Kathleen M. (2007). Maternal sensitivity and child responsiveness: Associations with social context, maternal characteristics, and child characteristics in a multivariate analysis. *Infancy*, *12*(2), 189-223.

Fernald, Anne, Marchman, Virginia A., & Hurtado, Nereyda. (2008). Input affects uptake: How early language experience influences processing efficiency and vocabulary learning. In Proceedings of the 7th IEEE International Conference on Development and Learning (pp. 37-42). IEEE.

Gleitman, Lila R., Cassidy, Kimberly, Nappa, Rebecca, Papafragou, Anna, & Trueswell, John C. (2005). Hard words. *Language learning and development*, *1*(1), 23-64.

Hoff-Ginsberg, Erika. (1994). Influences of mother and child on maternal talkativeness. *Discourse processes*, *18*(1), 105-117.

Huttenlocher, Janellen, Vasilyeva, Marina, Waterfall, Heidi R., Vevea, Jack L., & Hedges, Larry V. (2007). The varieties of speech to young children. *Developmental psychology*, *43*(5), 1062.

Imai, Mutsumi, Haryu, Etsuko, & Okada, Hiroyuki. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child development*, *76*(2), 340-355.

Imai, Mutsumi, Li, Lianjing, Haryu, Etsuko, Okada, Hiroyuki, Hirsh-Pasek, Kathy, Golinkoff, Roberta M., & Shigematsu, Jun. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child development*, *79*(4), 979-1000.

Landau, Barbara, & Gleitman, Lila R. (1985). *Language and experience* (Vol. 174). Cambridge, MA: Harvard University Press.

Newport, Elissa, Gleitman, Henry, & Gleitman, Lila R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. *Talking to children: Language input and acquisition*, 109-149.

Pan, Barbara A., Rowe, Meredith L., Singer, Judith D., & Snow, Catherine E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child development 76, 763*.

R Development Core Team. (2012). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Snow, Catherine E. (1972). Mothers' speech to children learning language. *Child development*, 549-565.

Snow, Catherine E. (1977). The development of conversation between mothers and babies. *Journal of child language*, *4*(01), 1-22.

Van Dijk, Teun A., & Kintsch, Walter. (1983). *Strategies in discourse comprehension*. New York: Academic Press.

Yuan, Sylvia, Fisher, Cynthia, & Snedeker, Jesse. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child development 83*, 1382-1399.