

# Second language acquisition of definiteness restrictions

Paul Hagstrom

Apr. 3, 2009

## 1 The L2A project

### Brief recap of arguments for UG from L1A

- Uniform stable state despite variation in experience
- Knowledge system is complex beyond evidence from the input
- Grammatical errors children make are few and confined
- Acquisition does not appear to involve conscious effort
- Biologically determined window of opportunity
- Language knowledge is largely opaque to introspection

*From these things we conclude...*

There is some kind of innate predisposition for language (acquisition) that governs the course of acquisition and the representation of language—“Universal Grammar” (“UG”).

### The role of UG in generative linguistics

Generative linguistics is essentially the exploration of the properties of UG. Current conceptions of UG take it to be comprised of:

**Principles** of representation, computation, and interface with language-external systems.

**Parameters** of variation between languages affecting particularly aspects of representation and computation.

**Acquisition** “device” of some kind that can determine the parameter “settings” from the language input.

UG defines a “space of possible languages,” but it is also part of the language knowledge itself, not *just* a “blueprint” for language knowledge.

## **Second language knowledge vs. native language knowledge**

The process of (adult) L2A seems quite different from L1A.

- Takes much longer.
- Requires a great deal of conscious effort.
- Rarely indistinguishable from native speech.

### ***A central question***

How much is knowledge of a second language like knowledge of a first language?

- Seems rather like learning calculus, or history.
- “Critical period”: (Acq. part of) UG loses effectiveness?
- Is second language knowledge really “language knowledge”?

## **Second language knowledge is systematic**

The knowledge of a second language that L2ers have does seem to be systematic. Perhaps on the basis of rules given during instruction or hypothesized during acquisition, but we can get meaningful results by asking second language speakers to judge sentences. And the L2ers will often show complex patterns in their judgments, like L1ers show in their native language.

One approach to determining how “language-like” second language knowledge is, we compare the subtle complexities of L2 knowledge and L1 knowledge.

## **Sources of L2 knowledge**

At least three sources from which complex L2 knowledge could arise:

**Instruction** being taught a complex rule, using it by rote.

**Transfer** building on the knowledge of the rules of the L1.

**UG** building on knowledge of “what languages are like” in a way that differs from the L1.

Is L2 knowledge different in kind from L1 knowledge? Is it completely dependent on L1 knowledge? Or is it “language-like” in and of itself?

The most interesting case is the last one—because it is the most surprising. This is where we’ll be focusing our energies.

## **Finding effects of UG in L2A**

The basic tactic here is:

- Identify a systematic pattern that is not explicitly taught...
- ...and differs from the L1...
- ...for which insufficient evidence exists in the input...
- ...but which is part of our idea of what languages are like...
- ...and look for evidence that L2ers have this knowledge.

## **2 The Definiteness Effect**

### **2.1 The DE in English**

#### **Acknowledgments**

This is ongoing work done in collaboration with Alyona Belikova, Öner Özçelik, and Lydia White (McGill University) and Tanja Kupisch (University of Hamburg). The main results reported here appear in White et al. (2009).

This research was funded by grants from the Social Sciences and Humanities Research Council of Canada (SSHRCC) and from the Fonds québécois de la recherche sur la société et la culture (FQRSC).

#### ***There*-constructions**

Many sentences like (1a) can be rephrased using *there*, as in (1b).

- (1) a. Several flies were in my soup.  
b. There were several flies in my soup.

There are some restrictions, however.

#### ***There*-constructions: Predicate type**

The predicate must be “stage-level”—that is, it must represent a temporary property.

- (2) a. Several flies were eating my soup.  
b. There were several flies eating my soup.
- (3) a. Several flies were unusually brave.  
b. \* There were several flies unusually brave.

### **There-constructions: Subject type**

One type of subject works (e.g., *several X*). Another type doesn't (e.g., *every X*).

- (4) a. Several flies were eating my soup.  
b. Every fly was eating my soup.
- (5) a. There were several flies eating my soup.  
b. \* There was every fly eating my soup.

### **The Definiteness Effect (“DE”)**

#### **Definiteness Restriction**

In a *there*-construction, the associate cannot be strong.

Definite NPs are in the second category (called “strong”). Indefinite NPs are in the first category (called “weak”). Here is a bigger list (Milsark 1977).

#### *Weak*

*a, some, many, few*, cardinal numbers (*one, two*), bare plurals (*seats, votes*), mass nouns (*cheese*)

#### *Strong*

*the, all, most, every*, demonstratives (*this, that*), possessives (*my, his*), personal pronouns (*I, you*), proper names (*Norm, Al*)

At least for the moment, it doesn't matter what exactly the property is that differentiates “strong” and “weak” in this context. Enç (1991) argues that it is actually a distinction of specificity, not of definiteness. Perhaps this might even differ between languages akin to the way Ionin (2006) proposes article systems vary, although this should not be the starting assumption. The main thing is: there is some property (which we assume is the same for Turkish and for English).

#### **Definiteness effect: Exceptions**

The subtlety: It is not as simple as ruling out “there is the...” however—this word sequence is acceptable in the right context.

- (6) *Deixis*:  
Oh, look. **There's the** missing soup ladle.
- (7) *List readings*:  
What should I read first?  
**There's the** article on determiners, and there's the *Harry Potter* book...

## 2.2 The DE in Turkish

### Turkish also shows the DE

We see definiteness effects in Turkish as well (Enç 1991)—it's not English-specific.

- (8) a. Bahçe-de birkaç çocuk var.  
garden-LOC some child exist  
'There are some children in the garden.'
- b. \*Bahçe-de o çocuk var.  
garden-LOC that child exist  
'There is that child in the garden.'

### Negative existentials in Turkish

But it is not precisely the same. Notably, in the *negative*, the DE disappears.

- (9) a. \*Bahçe-de o çocuk var.  
garden-LOC that child exist  
'There is that child in the garden.'
- b. Bahçe-de o çocuk yok.  
garden-LOC that child not-exist  
'There isn't that child in the garden.'

### Articles in Turkish

Another difference between English and Turkish is that Turkish does not have a definite article (Kornfilt 1997).

- (10) a. Ali kitab-ı okudu  
Ali book-ACC read  
'Ali read the book.'
- b. Ali bir kitab-ı okudu  
Ali a/one book-ACC read  
'Ali read a book.'
- c. Adam gel-di  
man come-PAST  
'The/a man came.'

## 3 Experimental investigation of the DE in L2A

### 3.1 Motivating the experiment

#### Poverty of the stimulus wrt DE?

**Transfer** The DE operates slightly differently in English and in Turkish. Specifically, in negative existentials definites are allowed in Turkish and disallowed in English. This allows us a way to detect L1 transfer.

**Instruction** The pattern is subtle—it involves not just *the* and *a*, but the various other weak and strong quantifiers. The pattern also has exceptions (list, deictic), so a simple rule—even if it were taught—would lead to non-English-like judgments.

**Frequency** It also can't be just that *There is the* is rare—it's only allowed in the exceptional contexts, and it is ultimately *more* frequent than the grammatical *There is one*.

#### Previous results on DE in L2 English (production)

- White (2003): Production data from one advanced endstate Turkish speaker.
- Lardiere (2005): Production data from one advanced endstate Mandarin speaker.
- White (2008): Elicited production (picture description) from 18 Turkish speakers and 15 Mandarin speakers, various proficiency levels

#### The results...

No DE violations were found.

#### Methodological questions about production data

- Production data can't distinguish between *ungrammatical* and *dispreferred*. But we want to know about grammaticality. (Some subjects didn't produce *there*-insertion at all. Should that really count as obeying the definiteness restriction?)
- In a production task, one has little control over the context.
- Certain sentence types did not occur in production data (i.e. negative existentials—the crucial test case for transfer).

#### To get at the questions more directly...

We designed a *judgment* task.

## 3.2 Subjects and materials

### Subjects and task

Subjects: Turkish-speaking learners of English. Proficiency: 3 low, 12 intermediate, 10 advanced. Tested in Istanbul.

Task: Contextualized acceptability judgments (n=90) on a computer screen, testing DE in affirmative and negative existentials, as well as deictic and list readings. (5 items per sub-category.)

### Sentence types: Affirmative existentials

- (11)
- a. Grammatical *there*-insertion with indefinites  
*There's a copy machine downstairs.*
  - b. Ungrammatical *there*-insertion with definites  
*There's the only copy machine downstairs.*
  - c. Grammatical *there*-insertion with weak determiners  
*There are several guests in the garden.*
  - d. Ungrammatical *there*-insertion with strong determiners  
*There are your keys on your desk.*

### Sentence types: Negative existentials

- (12)
- a. Grammatical *there*-insertion with indefinites  
*There doesn't seem to be a suitable bowl here.*
  - b. Ungrammatical *there*-insertion with definites  
*There isn't the bowl here.*
  - c. Grammatical *there*-insertion with weak determiners  
*There aren't many guests inside.*
  - d. Ungrammatical *there*-insertion with strong determiners  
*There aren't most guests inside.*

### Sentence types: Controls

- (13)
- a. Grammatical list contexts  
*Well, there's the bottle of wine from last night.*
  - b. Grammatical deictic contexts  
*Look. There's the mouse again.*
  - c. Grammatical indefinite subjects  
*A young woman is still waiting to see you.*
  - d. Ungrammatical indefinites  
*A kitchen stove is broken.*

### Sample test items

Grammatical *there*-insertion with indefinite article

**Lisa and Denis are moving into a new apartment. Lisa is worried about how to get their furniture upstairs. Denis says:**

**Luckily, there's an elevator in this building.**

*How natural is this sentence in this context? If you choose 'unnatural', please correct the sentence.*

- Natural  
 Not sure  
 Unnatural

### Sample test items

Ungrammatical *there*-insertion with definite article

**Anne is feeling sick, so she makes an appointment to see Dr. Salter. She arrives early and the nurse tells her to go right in, saying:**

**There's the doctor here already.**

*How natural is this sentence in this context? If you choose 'unnatural', please correct the sentence.*

- Natural  
 Not sure  
 Unnatural

### Sample test items

Grammatical *there*-insertion, weak DP, negative

**John was having a party. When Mary arrived, John suggested that she should join the others outside. He said:**

**There aren't many guests inside.**

*How natural is this sentence in this context? If you choose 'unnatural', please correct the sentence.*

- Natural  
 Not sure  
 Unnatural

### Sample test items

Ungrammatical *there*-insertion, strong DP, negative

**Some students have problems with an assignment, so they ask the secretary whether the statistics professor is available to help them. She says:**

**No, there isn't Professor Black in his office today.**

*How natural is this sentence in this context? If you choose 'unnatural', please correct the sentence.*

- Natural  
 Not sure  
 Unnatural

### Sample test items

Grammatical deictic with definite article

**Lucy and Jeremy saw a white mouse in their kitchen. They tried to catch it, but it escaped. When Lucy saw it later, she said to Jeremy:**

**Look, there's the mouse again.**

*How natural is this sentence in this context? If you choose 'unnatural', please correct the sentence.*

- Natural  
 Not sure  
 Unnatural

### Sample test items

Grammatical list reading with definite article

**Some diamonds were stolen from a jewelry store. The police asked who was in the store during the night. The manager replied:**

**There was only the security guard.**

*How natural is this sentence in this context? If you choose 'unnatural', please correct the sentence.*

- Natural  
 Not sure  
 Unnatural

### Coding: Relevant corrections

There's Mary at home.	⇒	Mary is at home.
There isn't the package here yet.	⇒	The package isn't here yet.
There aren't most guests outside.	⇒	There aren't many guests inside.
There's a suitable bowl over here.	⇒	That bowl over there should do the trick.

### Coding: Irrelevant/reverse

There was no one in class today.	⇒	There is no one in class today.
There's the bus at last.	⇒	At last there's the bus.
Ok. There's Mary at home.	⇒	Ok. There's Mary here.
There's a package for you at the post office.	⇒	Just to let you know, there's a package waiting for you at the post office.

### Coding: Discard

There's the only copy machine downstairs.	⇒	There's only copy machine downstairs.
Tony seems to be in the garden.	⇒	Is that Tony I hear singing in the garden?
There's another doctor here instead.	⇒	Dr. Smith is covering for Dr. Salter while she is away.
There's a suitable bowl over here.	⇒	That bowl over there should do the trick.

## 3.3 Experimental Results

**In brief:** There is a clear difference between grammatical and ungrammatical sentences, for both the intermediate and advanced L2ers: The L2ers know the DE constraint. Moreover, both groups of L2ers reject strong/definite NPs in negative existentials, even though these would be grammatical in the L1. They have not transferred the “Turkish DE” but are showing the DE as appropriate to English.<sup>1</sup>

---

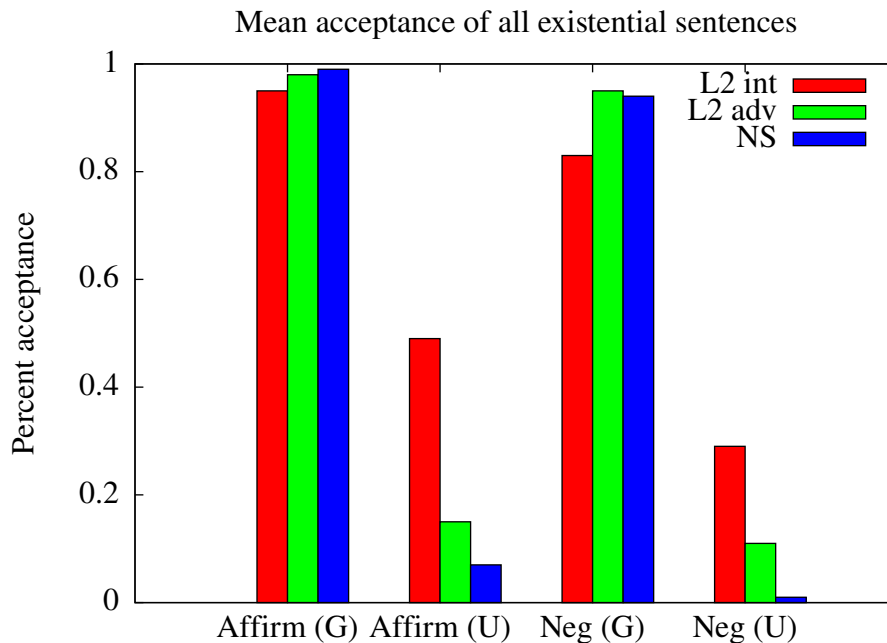
<sup>1</sup>Statistical information: Our test items included more ungrammatical existentials (definite articles, strong quantifiers and two other types of strong DPs) than grammatical (indefinite articles and weak quantifiers). To enable statistical comparisons, the analysis in Figure 1 is confined to articles and quantifiers.

We conducted 2 factor repeated measures ANOVAs to test for differences, then one factor ANOVAs and post-hoc Scheffé tests, as appropriate, to determine the source of any differences.

**Figure 1 (Overall comparison of existential sentences).** The 2-factor ANOVA shows a significant effect for group ( $f(2, 29)=6.044, p<0.01$ ), for sentence type ( $f(3, 29)=370.084, p<0.0001$ ), and a significant interaction ( $f(6, 29)=17.896, p<0.0001$ ). Post-hoc Scheffé tests show that the differences between grammatical and ungrammatical are significant for all groups, both for affirmatives and negatives. Comparing grammatical affirmatives and negatives, there are no significant differences for any group. For the ungrammatical sentences, the advanced group and native speakers did not differ significantly, but the intermediate group is significantly more likely to accept ungrammatical sentences than the advanced group and the native speakers (ungrammatical affirmatives:  $f(2, 29)=25.504, p<0.0001$ , ungrammatical negatives:  $f(2, 29)=14.218, p<0.0001$ ).

**Figure 2 (Ungrammatical affirmative existentials).** A 2 factor repeated measures ANOVA shows a significant effect for group ( $f(2, 29)=17.048, p<0.0001$ ), a significant effect for sentence type ( $f(3, 29)=5.896, p<0.001$ ) and a significant interaction ( $f(6, 29)=3.407, p<0.005$ ). The performance of the advanced subjects does not differ from the native speaker controls, whereas the intermediate subjects are significantly more likely than the controls and the advanced group to accept all ungrammatical existentials except for proper names. The intermediate group shows a significant difference across conditions ( $f(3, 11)=7.139, p<0.001$ ), with proper names being accepted significantly less

**Fig 1. Overall comparison of existential sentences**



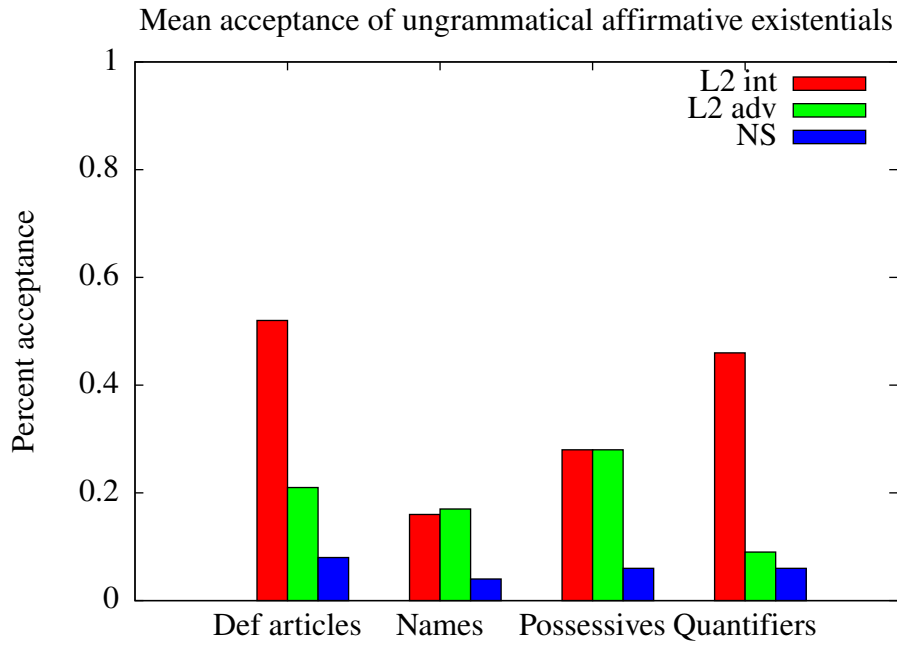
than definite articles and strong quantifiers. For the advanced group and the native speakers, there are no significant differences across conditions.

**Figure 3 (Ungrammatical negative existentials).** A 2 factor repeated measures ANOVA shows a significant effect for group ( $f(2, 29)=16.422, p<0.0001$ ), a significant effect for sentence type ( $f(3, 29)=7.807, p<0.001$ ) and a borderline significant interaction ( $f(6, 29)=2.204, p=0.05$ ). Native speakers reject all four sentence types outright, regardless of subtype, and the advanced L2ers show a similar pattern, with, again, no significant differences between them and the control group. The intermediate group differs from the native speakers on all four sentence types and they differ from the advanced group on possessives and quantifiers. Proper names are more consistently rejected than other ungrammatical cases. The intermediate group shows a significant difference across conditions ( $f(3, 11)=4.391, p<0.01$ ), with proper names being accepted significantly less than definite articles. The advanced group also shows a significant difference across conditions ( $f(3, 9)=50268, p<0.01$ ), with names and possessives both differing from definite articles. The native speakers show no significant difference across conditions.

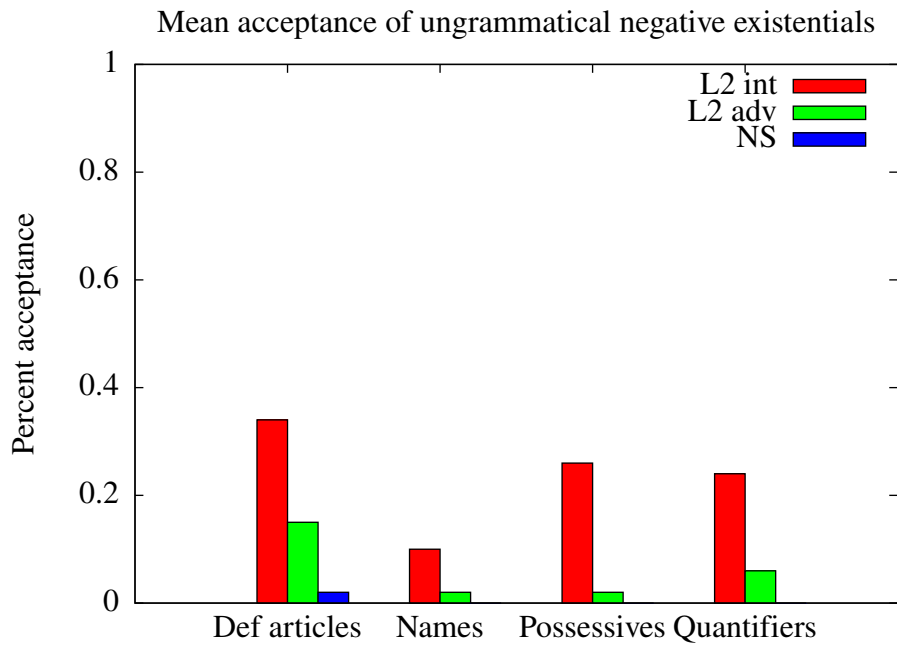
**Figure 4 (Control items).** In list (e.g. (7a)) and deictic (e.g. (7b)) contexts, it is possible for a definite expression to co-occur with there. A 2 factor repeated measures ANOVA (comparing list and deictic readings) shows a significant effect for group ( $f(2, 29)=6.695, p<0.005$ ), no effect for sentence type ( $f(1, 29)=2.116, p=0.1565$ ) and a borderline significant interaction ( $f(2, 29)=3.253, p=0.0531$ ). All subjects accept list readings as being natural. In other words, they are prepared to accept sentences involving there followed by a definite DP, which suggests that they are not simply following a strategy of rejecting all sentences involving definites, something which might otherwise account for their successful performance on the existentials. For the native speakers and the advanced group there is no significant difference between acceptances of list and deictic readings; however, the intermediate subjects are significantly less accepting of deictic readings ( $f(1, 11)=5.037, p<0.05$ ).

As for the ungrammatical sentences involving indefinite articles (such as (7c)), these were strongly rejected by the advanced group and the native speakers, and also rejected (to a somewhat lesser extent) by the intermediate group. These results suggest that the L2ers were not simply following a strategy of accepting all sentences involving indefinites, which, again, might otherwise account for their accuracy with existentials. A one factor ANOVA shows that there is a significant difference between the groups, with the intermediate group differing from both the other groups ( $f(2, 29)=7.236, p<0.005$ ).

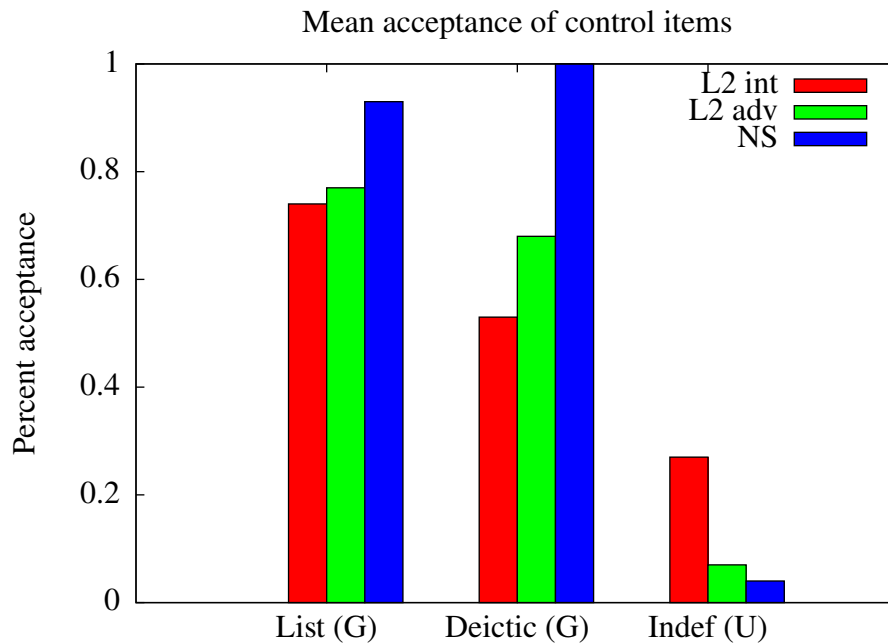
**Fig 2. Ungrammatical affirmative existentials**



**Fig 3. Ungrammatical negative existentials**



**Fig 4. Control items**



### Summary of results

- L2ers at all levels differentiate between grammatical and ungrammatical cases of *there*-insertion.
- Performance by the advanced group is not noticeably different from native speakers.
- No evidence of transfer: L2ers reject sentence types which are ungrammatical in the L2 and grammatical in the L1.

## 4 Discussion

### 4.1 Source of L2 knowledge?

#### Discussion

The L2ers' knowledge of English seems to include the English version of the DE. Where did they get it?

**Frequency?** No, L2ers show evidence of knowing how the DE applies, regardless of frequency of various determiners in this construction.

**Instruction?** (*There is a.../There are some...*). List readings (*There is the...*) were accepted.

**Transfer?** Not in an obvious way, negative existentials with definite DPs were rejected.

**UG?** Maybe. But what principles/parameters explain the DE?

## 4.2 Variation in the DE across languages

### Languages like Turkish

In Russian also, the DE does not hold in negative sentences.

- (14) \* V sadu est' vise deti.  
in garden exist all children  
(‘There are all the children in the garden.’)
- (15) V sadu net Ol'g-i  
in garden not.exist Olga-GEN  
(‘There isn’t Olga in the garden.’)

We believe that Japanese also behaves this way, maybe also Korean.

### Theoretical explanations for the DE?

One common view of indefinite noun phrases is that they are not themselves quantificational, but are bound by some kind of higher operator (Heim 1982; Kratzer 1998; Reinhart 1997). Diesing (1992) argued essentially that this binding can only apply to NPs in the deeper portion of the sentence structure, and that definite/strong NPs must be interpreted higher. In a *there*-construction, the associate is forced to remain within the indefinite domain, yielding ungrammaticality if the associate is strong. Though, this doesn’t really bring us a lot closer to understanding what’s going on.

### Speculations

One thing that Russian and Turkish have that differentiates them from English is that there is an actual (suppletive) form for ‘not-exist.’ Could this correlate with the ability to use definites in negative existential constructions? We don’t know. If so, maybe it points to a way for definites to escape the deep “indefinite part” of the tree in just those cases.

**Turkish** *var* ‘exist’, *yok* ‘not-exist’

**Russian** *byt* ‘exist’, *net* ‘not-exist’

**Korean?** *iss-* ‘be’, *eps-* ‘not-be’

## 4.3 Implications for L2A

### So what does it mean that the L2ers did so well?

- The question of transfer vs. UG is still somewhat open, pending an account of the DE.
- In its basic form, we have evidence *against* transfer—the DE in Turkish acts different from the DE in English, and the L2ers are using the English DE in English. So, if UG contains a parameter of language variation like “The-negative-DE parameter” (on for English, off for Turkish), then we have evidence that the L2ers were able to acquire a language system with the other setting for that parameter.

## How did they acquire it?

- What evidence led them to conclude that Turkish is not a neg-DE language? Is the morphological correlation a trigger?
- If so, if it's a syntactic difference, this reopens the UG vs. L1 transfer question—perhaps the DE-related constraints are invariant, but Turkish has a syntactic “escape hatch” allowing definites to escape the indefinite domain in the negative.
- Really, this just means we need to look in a different place (the syntactic positioning of arguments)—the basic distinction is still there, and still may only be derivable from UG.

## Representational issues

- One result: The L2ers are clearly distinguishing definite from indefinite (and strong from weak), given that they act on it. There does not seem to be any representation deficit (whether permanent or fluctuating). To varying degrees, contra Tsimplici & Mastropavlou (2007); Trenkic (2007); Ionin et al. (2004).
- Methodologically: even though in production L2ers often have problems producing articles (Huebner 1985; Thomas 1989; Robertson 2000), it is too hasty to take this as evidence that their representation or knowledge of these articles is impaired or underspecified.

## References

- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22(1): 1–25.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Ph.D. thesis, University of Massachusetts, Amherst.
- Huebner, Thom. 1985. System and variability in interlanguage syntax. *Language Learning* 35(2): 141–163.
- Ionin, Tania. 2006. This is definitely specific: Specificity and definiteness in article systems. *Natural Language Semantics* 14: 175–234.
- Ionin, Tania, Heejeong Ko & Kenneth Wexler. 2004. Article semantics in L2 acquisition: the role of specificity. *Language Acquisition* 12(1): 3–69.
- Kornfilt, Jaklin. 1997. *Turkish*. Routledge.
- Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide scope indefinites? In Susan Rothstein (ed.) *Events and grammar*. Dordrecht: Kluwer.
- Lardiere, Donna. 2005. On morphological competence. In Laurent Dekydtspotter, Rex A. Sprouse & Audrey Liljestrang (eds.) *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*, 178–192. Somerville, MA: Cascadilla Proceedings Project.
- Milsark, Gary. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3(1): 1–29.
- Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20: 335–397.

- Robertson, Daniel. 2000. Variability in the use of the English article system by Chinese learners of English. *Second Language Research* 16(2): 135–172.
- Thomas, Margaret. 1989. The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics* 10: 335–355.
- Trenkic, Danijela. 2007. Variability in L2 article production: beyond the representational deficit vs. processing constraints debate. *Second Language Research* 23(3): 289–327.
- Tsimpli, Ianthi-Maria & Maria Mastropavlou. 2007. Feature interpretability in L2 acquisition and SLI: Greek clitics and determiners. In Juana M. Liceras, Helmut Zobl & Helen Goodluck (eds.) *The role of formal features in second language acquisition*, 143–183. Erlbaum.
- White, Lydia. 2003. Fossilization in steady state L2 grammars: persistent problems with inflectional morphology grammars: persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6(2): 129–141.
- White, Lydia. 2008. Definiteness effects in the L2 English of Mandarin and Turkish speakers. In Harvey Chan, Heather Jacob & Enkeleida Kopia (eds.) *BUCLD 32: Proceedings of the 32nd annual Boston University Conference on Language Development*, 550–561. Somerville, MA: Cascadilla Press.
- White, Lydia, Alyona Belikova, Paul Hagstrom, Tanja Kupisch & Öner Özçelik. 2009. Restrictions on definiteness in L2 English. In Jane Chandlee, Michelle Franchini, Sandy Lord & Gudron-Marion Reiner (eds.) *BUCLD 33: Proceedings of the 33rd annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.