Measuring optionality in non-native grammars: A quantitative approach

Cristóbal Lozano
Universidad de Granada
http://www.ugr.es/~cristoballozano

- This presentation assesses syntactic optionality, a common phenomenon observed in second language (L2) acquisition that can be considered as a subtype of linguistic variability.
- Syntactic optionality is a pervasive phenomenon since it has been reported in different acquisition contexts (e.g., L1 acquisition, L2 acquisition, child bilingualism, L1 attrition).
- It is usual for L2 learners to optionally use two syntactic realizations (one of them being grammatical and the other ungrammatical) to express the same meaning.
- While optionality has attracted much attention in acquisition studies over the past decade, researchers have confusingly used the term "optimal" to refer to different types of variable learner behaviours which are sometimes unrelated (e.g., different and distinct types of divergence, indeterminacy and incompleteness). However, little is known about the precise quantitative properties that optional constructions in non-native, developing grammars must show in order to be considered truly optional.
- This study addresses optionality from a quantitative point of view and presents a statistical definition of optionality which will help future L2 researchers discriminate between truly optional L2 behaviour and other subtypes of variable (but not optional) behaviour.

What is optionality?
What does optionality look like?
  - Frequency?
  - Percentage?
  - Shape?
How can we discriminate between optionality and...
  - incompleteness?
  - divergence?
Is there a difference between native-like and near-native behaviour?

AIMS OF THIS TALK
- 'Optionality' is at centre of attention in SLR.
- BUT 'optionality' has been indiscriminately and confusingly used to describe different attainment patterns:
  - 'optional', 'variable', 'indeterminate', 'divergent'
- So, aim:
  - Describe optionality quantitatively.
  - Discriminate between optionality vs. other variable patterns in SLA.

STRUCTURE OF THIS TALK:
- Optionality: the phenomenon
- Optionality: qualitative definitions in the literature
- Optionality: source and explanations
- Optionality: quantitative properties
  - Criterion 1: type of comparisons
  - Criterion 2: instrument and measurement scales
  - Criterion 3: native categorical rules
  - Criterion 4: control group
- Detecting patterns:
  - Optionality here?
  - Near-native or native-like here?
- Conclusion

PRELIMINARY QUESTIONS

INTRODUCTION
THE PHENOMENON:
OPTIONALITY

GRAMMATICAL MANIFESTATIONS OF OPTIONALITY

L1 Fr – L2 Eng
- Alice quickly ate the mushrooms
- *Alice ate quickly the mushrooms

L1 Ger – L2 Eng
- For many kids living with their parents is a nightmare.
- *For many kids is living with their parents a nightmare

L1 Eng / Greek – L2 Spa
- A: ¿Quién vino anoche a la fiesta?
- B: Vino mi amiga
- *Mi amiga vino

OPTIONALITY: PERVERSIVENESS

Optionality is PERVERSIVE:
- Across linguistic levels and structures:
  - Morphology
  - Syntax
- Across acquisition contexts:
  - Adult L2 acquisition
  - Child L1 acquisition
  - Child bilingualism
  - Child SLI
  - L1 attrition
- Across developmental stages:
  - Initial states
  - Intermediate states
  - Advanced states
  - End states
- Across different L1 – L2 combinations
  - English, Spanish, Italian, Greek, German, Chinese, Hungarian...

For overviews, see all studies by Sorace and references therein, as well as refs at the end.
English native dative alternation: Optionality here?
- I gave her a book
- I gave a book to her

Both options are grammatical, BUT...
- A: What did you give to Sue?
- B: I gave her a book about applied psycholinguistics
  I gave a book about applied psycholinguistics to her

And now?
- I gave a man who was begging in the street a dollar
- I gave a dollar to a man who was begging in the street

Principle of end-focus and Principle of end-weight (Arnold et al. 2000)

Principle of non-synonymy of grammatical forms:
- A difference in syntactic form spells out a difference in meaning and/or pragmatic function (Bolinger 1977)

So, optionality in native grammars can be ‘apparent’, a.k.a. ‘pseudo-optionality’ (Papp 2000)

‘Apparent’ native optionality ≠ L2 optionality
- Native: each variant has a different meaning/interpretation
- Non-native: both variants are in free variation and have the same meaning/interpretation

**Optionality in Native Grammars?**

**Sorace:** pioneering work on syntactic optionality
- Sorace (1993, 2000 and other works in references)
- Sorace (2000): coexistence in the learner’s interlanguage of two phonological forms for one logical form.

Let 1, and 12 typically (but may not) make use of the same lexical resources.
- John often eats pasta
  *John eats often paste
- I draw a triangle
  *I draw Ø triangle

**Definitions of Optionality**

« The term ‘optionality’ refers to the use of two or more competing forms with the same meaning. [...] it refers to the variable use and nonuse of a particular surface feature, with no difference in meaning between the two forms. »
- Robertson (2000:138)

Optionality is the « simultaneous presence in a learner’s grammar of two features that should be mutually exclusive »
- Truscott (2006:311)

« L2 grammars clearly have a much greater degree of tolerance for synonymy than native grammars. »
- Sorace (2000:98)
- BUT cf. Bolinger’s Principle of non-synonymy in native grammars bans optionality
Optionality in generative grammar: performance phenomenon
- Performance is out of scope of gen. grammar

Minimalism (Chomsky 1995 and later work)
- Grammatical representation of features:
  - Discrete value: either [+1] or [-1] value
  - Mutually exclusive
  - Once a value is correctly acquired/set, learner should not ‘go back’ to incorrect value.

BUT: optionality and variability well known for decades in SLA
- Well attested phenomenon in syntax and morphology
- Pervasive: surfaces in wide range of constructions
- Backsliding (Selinker 1972): after setting correct value, learners do often ‘go back’ to incorrect value.

EXPLAINING OPTIONALITY

OPTIONALITY
AND LING THEORY

INPUT: Ambiguity
COMPETENCE: Feature deficits
PERFORMANCE: Processing defects

INPUT:
- Processing deficits
- Interfaces
- Impairment of features

COMPETENCE:
- Feature deficits
- Competing forms (same grammar)
- Blocking failure

PERFORMANCE:
- Development
- Mapping of features
- Input

SOURCES OF OPTIONALITY

SOURCES OF OPTIONALITY: A SUMMARY

Van Kampen (2004)
Sorace (2006)
Prévost & White (2000)
Parodi & Tsimpli (2005)
Domínguez & Arche (2008)
Papp (2000)
Diachrony: Zobl & Liceras (2005)
Diachrony: Zobl & Liceras (2005)
Defining optionality from a **linguistic** point of view is essential.
- We need to know its **ling** properties.
- We need to know its **causes(s)**.
- We need to know the **stages** when it should appear.

BUT defining optionality from a **quantitative** point of view is crucial.
- We need to distinguish between optionality and other **variable patterns** in SLA.
- If we can distinguish between **optional pattern** vs. other **variable patterns**, we can then **redefine** and **refine** optionality theoretically.

**THE STORY SO FAR**

- **Quantitative properties of optionality**
  - **Overlooked** area in the literature.

    - Addresses issue of **theoretical aspects** of optionality.
    - More interested in **pervasiveness** of optionality (L2, L1, SLI, attrition).
    - Focus on **explaining optionality** (deficits at interfaces or deficits when processing).

  - **Rest of authors**
    - Interested in **explaining optionality** (defective features, competing grammars, deficits with processing ...)

  - **Rest of studies:**
    - **PROBLEM:** they claim they discuss ‘optionality’, but they use the term **confusingly** to refer to **different patterns**.

**THEORETICAL LIT:**

**QUANTIFYING OPTIONALITY**

**BASIC CRITERIA FOR MEASURING OPTIONALITY**
**POSSIBLE PATTERNS OF L2 VARIABILITY**

- **Sorace (1993)**
  - **Convergent**: when learner representation is native-like.
  - **Divergent**: when learner representation consistently diverges and differs from the natives’.
  - **Optionality** is a subtype here.
  - **Incomplete**: when learners show indeterminate intuitions.

- **Papp (2000)**
  - Working definition of optionality (see next chart)

**POSSIBLE PATTERNS OF ATTAINMENT**

- **Convergent**
- **Divergent**
- **Incomplete**

**Basic criteria for measuring optionality:**

**CRITERION 0: COMPARISONS**
In the L2 literature: only within-group comparisons

- If grammatical sentence is preferred significantly more than the ungrammatical sentence, then native-like competence.
- Example:

<table>
<thead>
<tr>
<th></th>
<th>Grító una mujer</th>
<th>*Una mujer gritó</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natives</td>
<td>9</td>
<td>&gt; 5</td>
</tr>
<tr>
<td>Learners</td>
<td>9</td>
<td>&gt; 7.5</td>
</tr>
</tbody>
</table>

**Comparisons:**

**Within-group and between-groups**

White (2002:26-27)

- Trying to compare learners’ performance on a given type of construction against natives’ performance on the same construction, leads to ‘comparative fallacy’.
- She argues that the crucial point is learners distinguish between grammatical vs. ungrammatical, as natives do:
  - «If certain sentence types are treated significantly differently from other sentence types, this suggests that the interlanguage grammar represents the relevant distinction (whatever it may be), even if the degree to which L2 learners observe it in performance differs from that of native speakers.» (p. 26-27)

Parodi & Tsimpli (2005:267)

- «Since we are interested in the nature and degree of optionality in L2 grammars we compare the rankings for the different structures tested within a group (and not the absolute ranking across groups).»

Lozano (2006a)

- Preliminary approach to L2 acquisition patterns.
- Statistical definition of optionality.
- Within-group and between-group analyses are necessary.

**Crucial:**

- IF we want to clearly discriminate between different L2 behaviour patterns,
- THEN both within-group AND between-group comparisons.
- WHY? See next charts and rest of charts in presentation.

Which pattern here?

- Clearly:
  - **Native-like pattern**
- But what about next scenario, nº 6?

Additional scenario, nº 5
- Native-like or near-native?
- If between-group analysis
  - then native-like
- If both between-group AND within-group
  - then near-native

NEAR-NATIVE PATTERN
SCENARIO Nº 6

- Near-native or optional?
- Clearly optional

OPTIONAL PATTERN
SCENARIO Nº 7

- Comparisons must be done
  - Within groups
  - Between groups

- They can be measured with two-way Anova with Post-hoc tests or with simple t-tests
  - At least 2 independent variables with 2 levels:
    - Group [native/learner]
    - Construction [grammatical/ungrammatical]

- Comparisons at α=0.05 level of statistical significance.

CONCLUSION
ON COMPARISONS
Studies on optionality use a wide variety of instruments

- **Production data:**
  - Elicited written production (Sopata 2005, Martínez Adrián 2007)
- **Comprehension data:**
  - Acceptability judgements (Papp 2000)
- **Behavioural data:**
  - Reaction time (Beck 1998)

**Measurement scales**

- **Likert scales:**
  - 10-point positive scale (Papp 2000)
  - Positive-negative scale (Lozano 2006a)
- **Other scales:**
  - Frequency corpus data (Osborne 2006, Robertson 2000)

**Different instruments** will yield different rates of production/acceptance.
- This is normal.
- **Different measurement scales** and frequency data will yield different behaviour patterns, **BUT**...
- If scales are **STANDARDIZED**, then results will be comparable.
- then optionality can be detected more accurately.
- **Standard Likert-type scale:**
  - Positive: 1 2 3 4 5
  - Negative: -2 -1 0 1 2
- **Raw frequency data (corpus):**
  - 7/14 (i.e., 7 articles produced out of 14 contexts)
Then, results can be standardized
- Either shown in %
- Or shown from 1-10
**Note:** all the charts in this presentation have been standardized from different measurement scales and frequency data, so they are all comparable.

**Why Likert-type scales?**

- They can show the full range of optionality as opposed to dichotomous yes/no scales.
- Also, both variants (grammatical vs. ungrammatical) need to be compared so as to observe different patterns:

  - **Native-like:**
    - Una mujer gritó
    - Gritó una mujer

  - **Indeterminate:**
    - Una mujer gritó
    - Gritó una mujer

  - **Optional:**
    - Una mujer gritó
    - Gritó una mujer

**Positive-negative Likert scale**

- Acceptability rate
  - Categorical: Convergent, Divergent
  - Non-native: Convergent, Divergent, Incomplete

**Indeterminacy is also clearly appreciated**

- Rejection rates can be clearly appreciated
Yuan (2001):
- To obtain potential data concerning optionality of thematic-verb raising in L2 Chinese, the subjects were told that they could mark both sentences as correct if they thought that was the case. Conversely, they could also mark both as incorrect if they believed the two sentences were wrong. There was also a choice of ‘I don’t know.’ (p. 260)
- [original sentences in Chinese]
  - a. My brother drinks usually German wine □
  - b. My brother usually drinks German wine □
  - c. I don’t know. □
- BUT: can we really measure optionality with Yuan’s instrument?

Robertson (2000)
- Controlled production (corpus) data: ok
- Count frequencies of supplance of full form (the) vs null form (0) in obligatory occasions.
- Then transform frequency data (count) to percentage (%)
  → OK.

Sopata (2005)
- Grammaticality judgement.
  - *Auf der Autobahn wird es sehr schnell gefahren. on the highway is it very fast gone
  - Correct: Sure □ Quite sure □ Quite unsure □
  - Incorrect: Sure □ Quite sure □ Quite unsure □
  - Don’t know □
- But can we detect optionality with this scale?

Basic criteria for measuring optionality:

**CRITERION 3: NATIVE CATEGORICAL RULES**

To investigate L2 optionality, we need a clear native categorical rule.
- John often eats pasta
- *John eats often pasta

If we use a native pseudo-optional or even native optional rule, we will never be sure whether L2 optionality is result of...
- Optionality in the input
- Developmental issue
  - I gave a book to Sue
  - I gave Sue a book
"...since input and output converge in this case, it is unclear whether the optionality observed is a property of the developing grammars or, simply, target-like performance" (p. 280).

Of course, native optional constructions should be ruled out in the study of non-native optionality.

The results for the native Greek speakers show a significant preference for a clitic \( [p<0.001] \) ... However, the learners' judgements do not show a clear preference for PC [purposive clauses] with or without a clitic, although the upper group of learners show the tendency to approximate native-speaker behaviour. The lower group gives a picture of 'real optionality' (n.s. values)." (p. 269)

Though use of control group seems obvious, some studies fail to do so.

- **Control group is necessary:**
  - To ensure we have only native categorical rules.
  - To ensure that ling theory is correct.
  - To ensure between-group comparisons with learners.
**Domínguez & Arche (2008)**

- The optional result in natives is unexpected, since differences have been reported (Lozano 2006a, 2006b) for natives.
- Then, EITHER theory is wrong and their data are correct OR their method is wrong and theory is correct.

**AIM**

- To show that some researchers discuss “optional” behaviour, but they refer to different patterns:
  - Truly optional
  - BUT also near-native (close to natives, but not yet)
  - And also indeterminate.

**THE QUESTION:**

- Are the following patterns truly optional?

---

**Osborne (2008)**

"This is in line with what would be expected in a case of optionality - that post-intermediate learners show a pattern of usage which tends overall towards that of native speakers, but which diverges in varying degrees..." (Osborne 2008).

So, optionality is ‘divergence’!! But, according to my analysis, we have a near-native pattern here.
Let's transform the scale.

As for non-finite forms, Prévost & White (2000:118) conclude that "there is variability (=optionality) in verb placement. With the exception of Abdelmalek, ... The other subjects show non-finite verb forms in raised and unraised positions."

The "both correct" option is treated as a reflection of optionality: "In other judgments (between 17% in the third and 35% in the second group) both sentences are accepted, which indicates clearly a non-target-like optionality in the use of expletives claueded internally." (Sopata 2005:189)

BUT do these values (18%, 17% ...) represent true optionality?

The % of participants who gave a particular response is also considered as a valid indicator of optionality: "(both correct" option in yellow): "...over half of the participants from groups 2, 3 and 4 accept the optional appearance of expletives...". (Sopata 2005:189).

BUT this simply indicates that around half of the participants in each group show optionality, and not that each group of participants accepts the grammatical AND ungrammatical sentence simultaneously.
NEAR-NATIVE OR NATIVE-LIKE?

AIM
- To show that some researchers discuss “native-like” behaviour, but they refer to different patterns:
  - Native-like (just exactly like natives)
  - Near-native (close to natives, but not yet)
  - Incomplete (more or less indeterminate)

THE QUESTION:
- Are the following patterns native-like or near-native?

**Domínguez & Arche (2008)**

![Graph showing acceptability rate in different groups](image)

*The UG [advanced] group was the only group to behave like the native speakers, since both the intermediate and the beginner groups showed the reverse behaviour accepting the non-inverted structures and rejecting the inverted ones.* (p. 102)

BUT advanced do not show native like behaviour, rather indeterminate or, at most, optional!!

**Sorace (1993):**

![Graph showing acceptability rate in different groups](image)

*The judgements of the French subjects on sentences exhibiting clitic-climbing are similar to the native judgements.* (p. 40)

Sorace only performs within-group comparisons.
**CONCLUSIONS**

- Different attainment patterns need a solid...
  - **Theoretical** description
  - **Quantitative** description

- Jansen (2000)
  - Analysed a series of key SLA studies on the UG-access debate. They all used the same data, but obtained different results and conclusions.
  - «...when it comes to underpinning theoretical claims, more rigour in data description should be exercised. » (p. 27)
  - «If researchers wish to pursue their quest for the principles underlying the process of second language acquisition, they need good tools – predictive, testable theories and rigorously analysed data – in order to test their theoretical claims. » (p. 41)

- The term ‘optionality’ has been abused in SLA research, but we need more fine-grained measure of it, because...
  - We cannot understand the CAUSES of optionality UNLESS we have the tools to **DETECT** it and **DISCRIMINATE** it from other behaviour patterns.