Language Faculty Science 3

- Replication in LFS
  - Including:
    - On-line “Experiments” in Japanese
    - *Off-line* Experiments in Mandarin Chinese, Korean, and English

- Among the general conceptual points about LFS
  - “Factual knowledge” vs. “Comprehension” (The latter involves deduction of predictions, as in our “... → BVA(X, Y):no”, while the former does not.)
  - The general significance of success of LFS

Replication in LFS

In our “Experiments”, we had only two choices of X and two choices of Y for BVA(X, Y).

X: every N, more than one N
Y: his, their

Revised Hypothesis (stated slightly differently here):
BVA/DR/Coref(X, Y) is possible only if X c-commands Y, provided that X is [-XQrk] and Y is [-YQrk].

BVA(X, Y)

BVA(**every** N, his)
BVA(**every** N, their)
BVA(**more than one** N, his)
BVA(**more than one** N, their)

DR(X, beta)

DR(**every** N, three/four Ns)
DR(**more than one** N, three/four Ns)

Coref(alpha, Y)

Coref(John, his)
Coref(John, their)

When the LFStist (language faculty scientist) conducts an experiment on themselves, they consider many more choices of X and many more choices of Y.

More than 20 choices of X and more than 10 choices of Y, for example.
Instead of 2x2=4 combinations of X and Y as in our “Experiments”, more than 20x10=200 combinations of X and Y are routinely checked.

Our "Experiments" (= the on-line "Experiments" mentioned in First and Second Lectures on LFS) only contain the following four schemata (=sentence patterns), focusing on cases where Y precedes X.

Y's N, X Verb (e.g., *His father, every boy praised*, *His father, John praised*)
Y's N Verb X (e.g., *His father praised every boy*, *His father praised John*)
Y, X Verb (e.g., *Three students, every teacher praised*)
Y Verb X (e.g., *Three students praised every teacher*)

The *basic* self-experiment by an LFStist considers the following 16 basic schemata (sentence patterns), and checks the possibility of BVA(X, Y), DR(X, Y) and Coref(X, Y) for each schema.
(The ones included in our "Experiments" are in red font.)

(1) Eight basic schemata in the case of English, where Y precedes X)

1	Y's N, X Verb
2	Y's N Verb X
3	Y, X Verb
4	Y Verb X
5	Y's N, X's N Verb
6	Y's N Verb X's N
7	Y, X's N Verb
8	Y Verb X's N

(2) Eight other basic schemata in the case of English, where X precedes Y)

1	X Verb Y's N
2	X, Y's N Verb
3	X Verb Y
4	X, Y Verb
5	X's N Verb Y's N
6	X's N, Y's N Verb
7	X's N Verb Y
8	X's N, Y Verb

In the *basic* self-experiments by Japanese LFStists, they routinely deal with four distinct "verb types", instead of just one.

16x4 =64 schemata are therefore routinely checked, along with various combinations of X and Y.

Note: This does not include the checking of effects of other factors that cannot be discussed in this course.

The *basic* correlational prediction discussed (the prediction that no Reds in the middle intersection) has been supported by results of self-experiments by the LFStist, over and over (although what choices of X and Y lead to Green in the middle intersection *at a given time* can vary).

That already constitutes replication within self-experiment by the LFStist across different choices of X and Y, across different verb types, and across different schemata (sentence patterns).

The LFStist then examines new, sometimes quite complicated, schemata (sentence patterns) and checks a number of new predictions under new hypotheses about those new schemata (and related issues), by using the combinations of X and Y that are effective for checking c-command effects *at a given time*, to find out about what merge-generated sets those new schemata correspond to, thus expanding their empirical coverage.

Continuation of this slide is given at the end of this file.

That is about the self-experiment.
What about non-self experiments?

On-line “Experiments” in Japanese.

Just like our “Experiments”, a replicational demonstration attempt dealing with Japanese contains much fewer choices of X and Y (and much fewer choices of schemata) than the LFStist checks in their self-experiments.

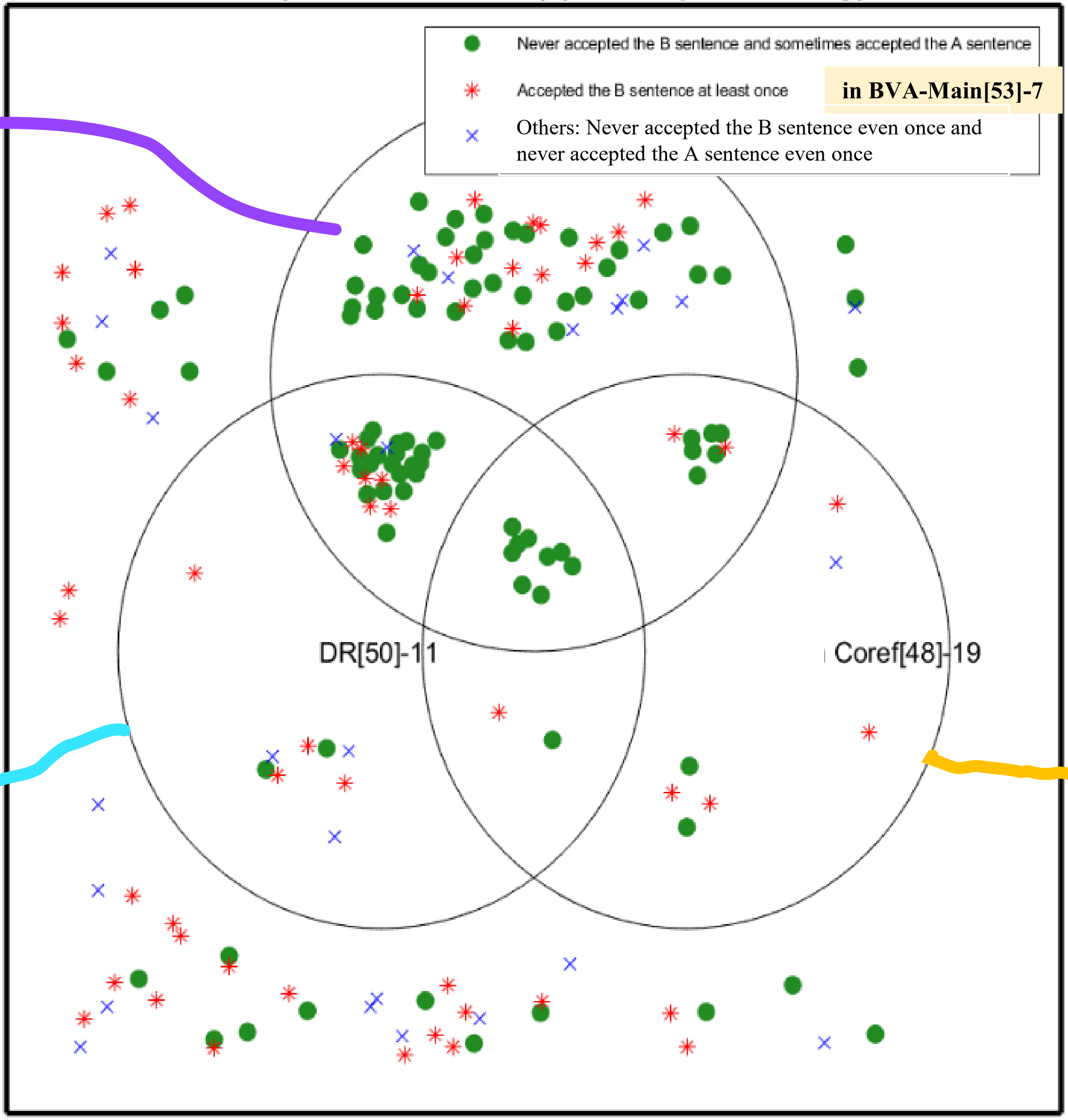
In one such attempt (about 200 undergraduate students in an GE course in a Japanese university), there are five choices of X and two choices of Y.

The results of replicational demonstration attempts support the correlational prediction: If someone is in the middle intersection, that person is not a Red *.

Just one example as an illustration.

BVA53-7

Those who passed just DR-Lex-Sub49-3 test.



The numbers after "BVA", "DR", etc. are Experiment #'s in on-line "Experiments" in Japanese.

Green with DR(X, beta) (Those who consistently rejected the Schema B sentences and accepted the Schema A sentences at least 25% of the time in DR50-11.)

Green with Coref(alpha, Y) (Those who consistently rejected the Schema B sentences and accepted the Schema A sentences at least 25% of the time in Coref48-19.)

In *every other* combination of X and Y for BVA(X, Y), there are no Reds in the middle intersection; furthermore, there are some Greens there.

This provides us with replicational demonstration of c-command detection.

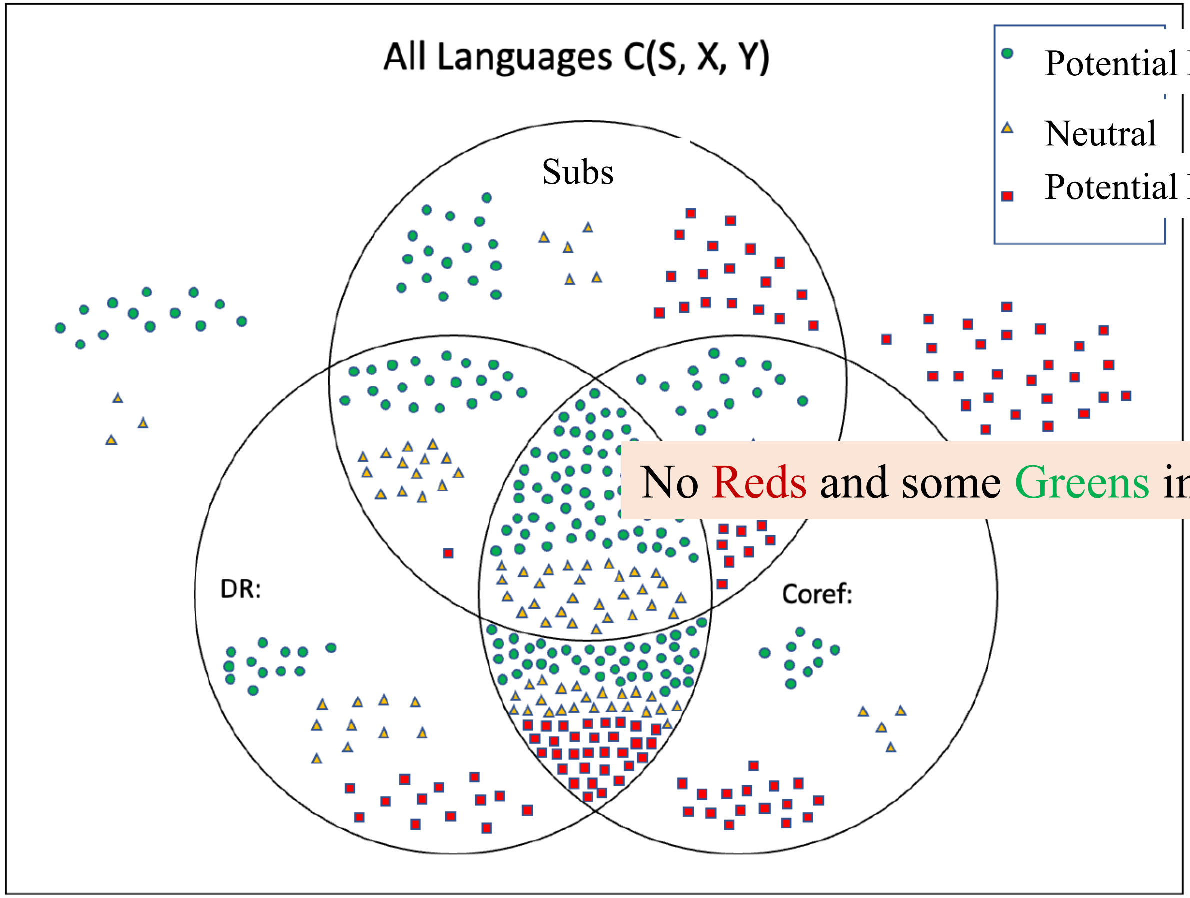
This can be taken as strong evidence for c-command effects and hence for the existence of the Computational System of the language faculty, containing Merge, as suggested by Chomsky.

Chapter 6: Replication: Predicted Correlations of Judgments in Japanese contains the details.

How about other languages?

Daniel Plesniak has worked on English, Mandarin Chinese and Korean.

(369) a. Venn Diagram



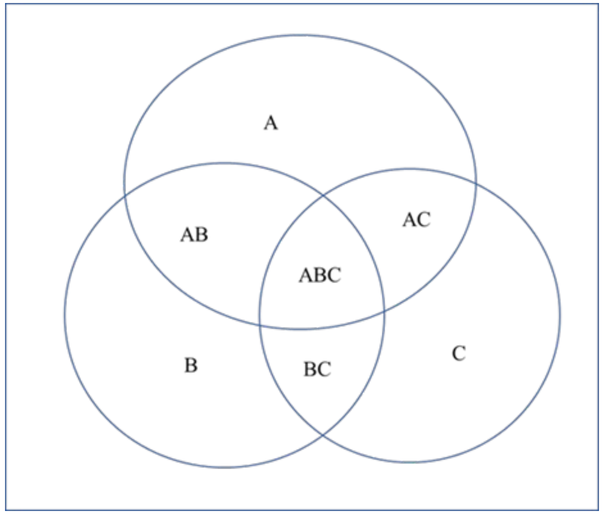
No Reds and some Greens in the middle.

Results based on off-line "Experiments" in English, Mandarin Chinese, and Korean. (12, 10, and 11 speakers, respectively. Each "Experiment" was one-on-one and lasted about an hour.)

Each dot represents a "data point" (a set of judgments on sentences with particular choices of X, and Y of BVA(X, Y)), not a speaker. Where each dot belongs is as we have discussed in class.

b. Summary table

# of Results	Green	Yellow	Red	Total
ABC	66	29	0	95
AB	20	14	1	35
AC	15	6	8	29
BC	41	23	31	95
A	16	5	18	39
B	12	10	13	35
C	9	4	16	29
None	10	3	26	39
Total	189	94	113	396

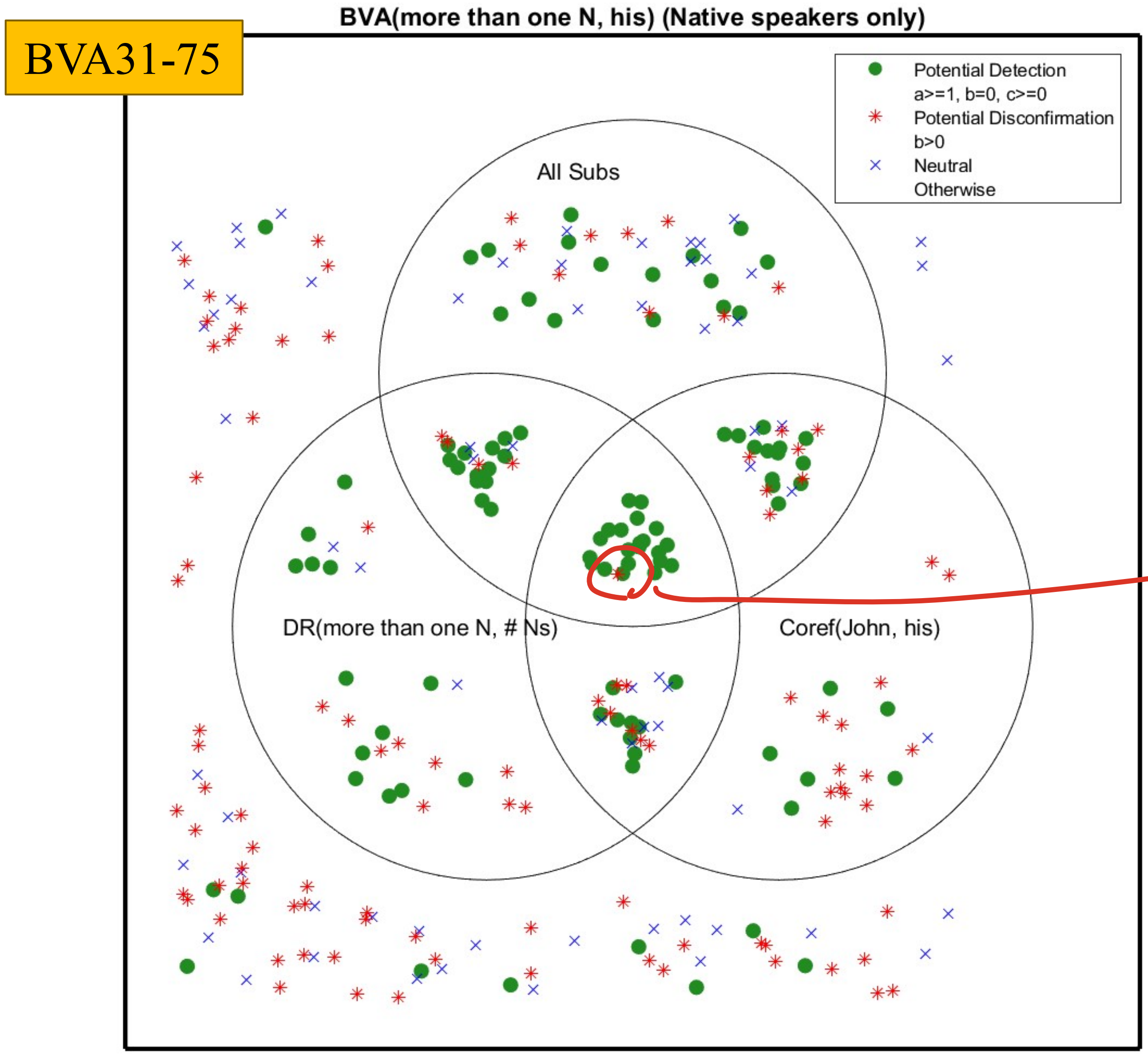


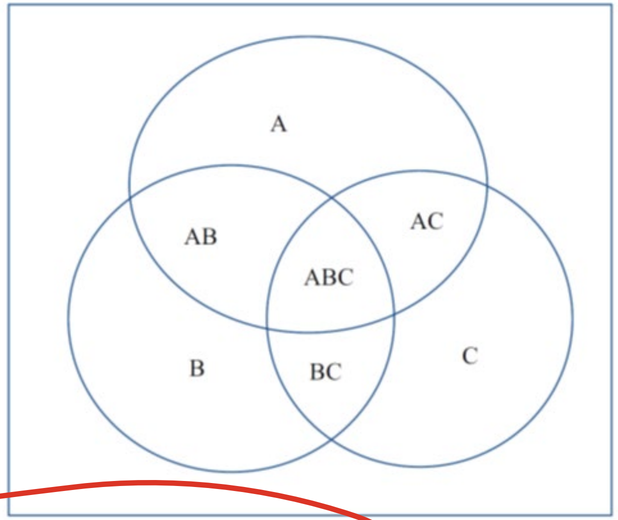
Adapted from Plesniak 2022: (369) on page 440.
USC Dissertation: Towards a Correlational Law of Language: Three Factors Constraining Judgement Varation

Results of these demonstration attempts should, however, be understood with a degree of caution, given the very limited choices of X and Y and other factors, as compared to what is checked in the LFStist's self-experiments.

This applies to results that are in line with our predictions, such as those mentioned, as well as to results that are not in line with our predictions, such as the result indicated in the following slide, discussed in the second lecture.

BVA(more than one N, his) BVA31-75 (Green, Red, and Blue here are about BVA31-75!)
Circle B: DR(more than one N, # Ns) Potential Greens in DR63-19
Circle C: Coref(Johh, his) Potential Greens Coref82-10
Circle A: Passing all the Subs with “more than one N” (not with “every N”) (31-77, 31-76, 63-13), and 63-10 and 31-48





Results	Green	Blue	Red	Total
ABC	21	0	1	22
AB	14	3	4	21
AC	13	4	7	24
BC	9	7	7	23
A	16	15	9	40
B	13	3	10	26
C	6	2	14	22
None	10	35	55	100
Total	102	69	107	278

If we change the %Value(STA) from “higher than 1” (the percentage of the Yes answers among all the answers given on Schema B is 1% or higher) to “100” (all the answers given on Schema B are “Yes”), the result becomes like this:

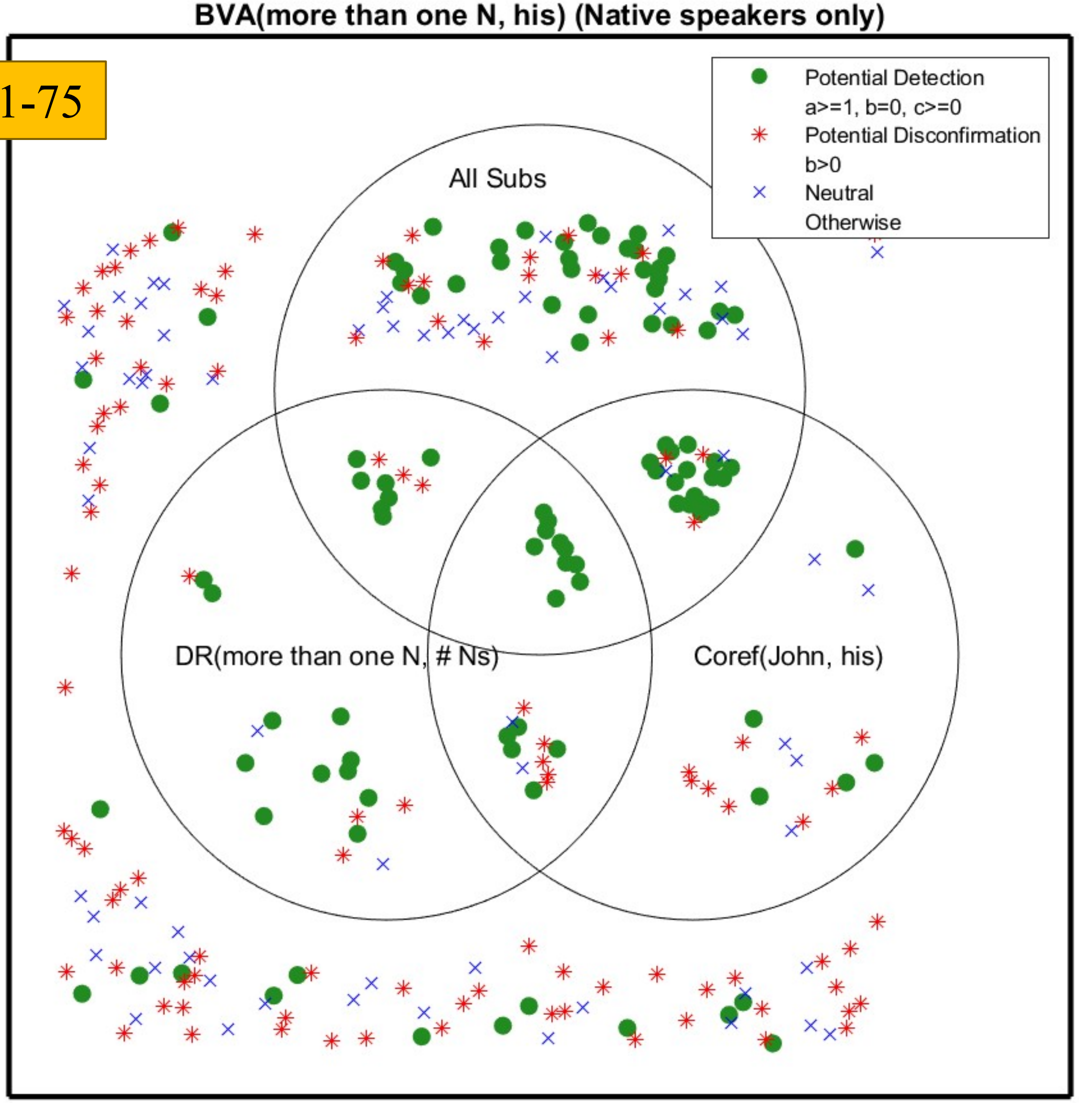
BVA(more than one N, his) BVA31-75 Green, Red, and Blue are determined as before

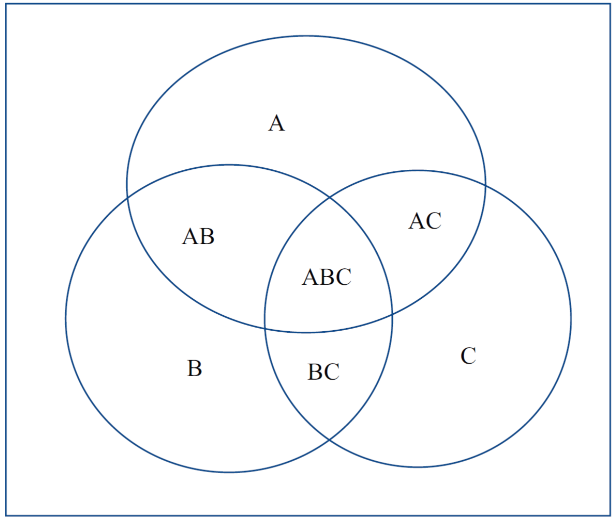
Circle B: DR(more than one N, # Ns) Green in DR63-19, based on %Value(STA)=100 and %Value(STB)=0

Circle C: Coref(Johh, his) Green in Coref82-10, based on %Value(STA)=100 and %Value(STB)=0

Circle A: Passing all the Subs with "more than one N" (not "every N") (31-77, 31-76, 63-13), and 63-10 and 31-48 as before

BVA31-75





Results	Green	Blue	Red	Total
ABC	10	0	0	10
AB	7	0	3	10
AC	17	2	3	22
BC	5	2	5	12
A	30	20	15	65
B	11	2	4	17
C	5	5	8	18
None	17	38	69	124
Total	102	69	107	278

We see Reds in the middle intersection only very occasionally, especially when we use very stringent criteria for inclusion in Circle A, Circle B, and Circle C.

If we conduct demonstration attempts (such as our on-line “Experiments” and their Japanese counterparts) infinitely many times, with infinitely many speakers, however, we are bound to see some Reds in the middle intersection!!

Why?

Because participants in demonstration attempts are not nearly as reliable as the LFStist.

The LFStist tries to check their judgments across BVA, DR and Coref very closely, by going back and forth among them, by going back and forth among different choices of X and Y, among different Schemata (different sentences patterns), and among other variables not considered above.

By checking things back and forth like that, LFStist can try to avoid effects of "change" of their "cognitive state" that might result in a shift from [-YQrk] to [+YQrk] for a given choice of Y, which can indeed happen (and other shifts like that).

Participants in demonstration attempts do not (know how to) do that.

When they are allowed to report their judgments over a long period of time, as in the demonstration attempt in Japanese (where students participated in on-line "Experiments" throughout the entire semester), they may be reporting their judgments based on different "cognitive states", which might result in shifts of the type just noted, thereby making the intended correlation of judgments not showing up as predicted.

They can simply make a mistake pressing a radio button. (The current design of the on-line "Experiments" does not allow the participant to self-correct their "errors"; this can potentially result in there being a Red * in the middle intersection by an error!)

So, Green and Red in demonstration attempts are not as reliable as Green and Red in the LFStist's self-experiments.

The fact that we seldom get Reds in the middle intersection in demonstration attempts indeed makes us hopeful that our hypotheses and the general correlational methodology in LFS may be on the right track.

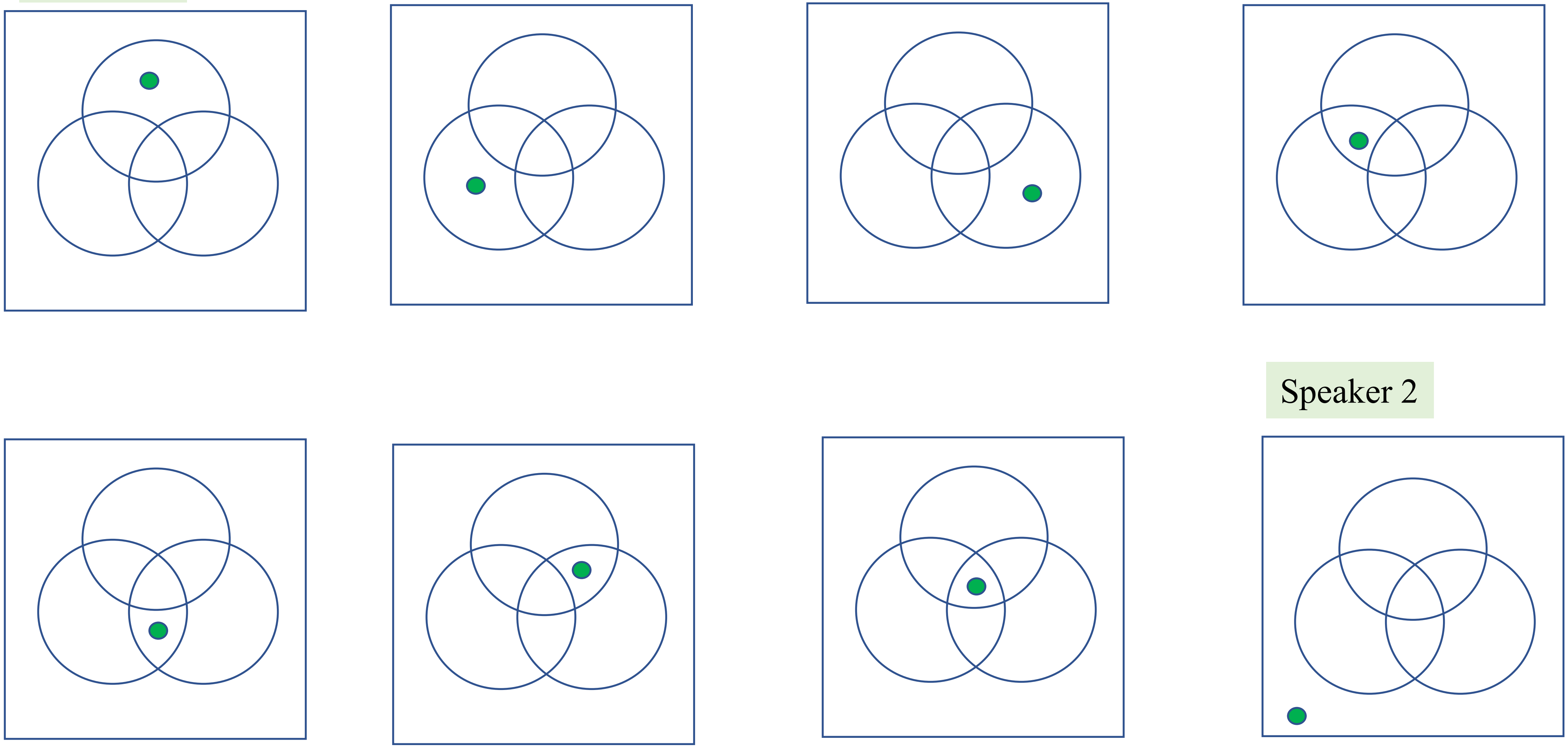
The language faculty is internal to an individual.

Our predictions are, therefore, about an individual.

For each individual speaker, we have one of the following 24 Venn Diagrams, as noted in Second Lecture of Language Faculty Science. (8 possibilities about where the speaker is, and 3 possibilities of the speaker's color (Green, Red, or Blue). 8x3=24.

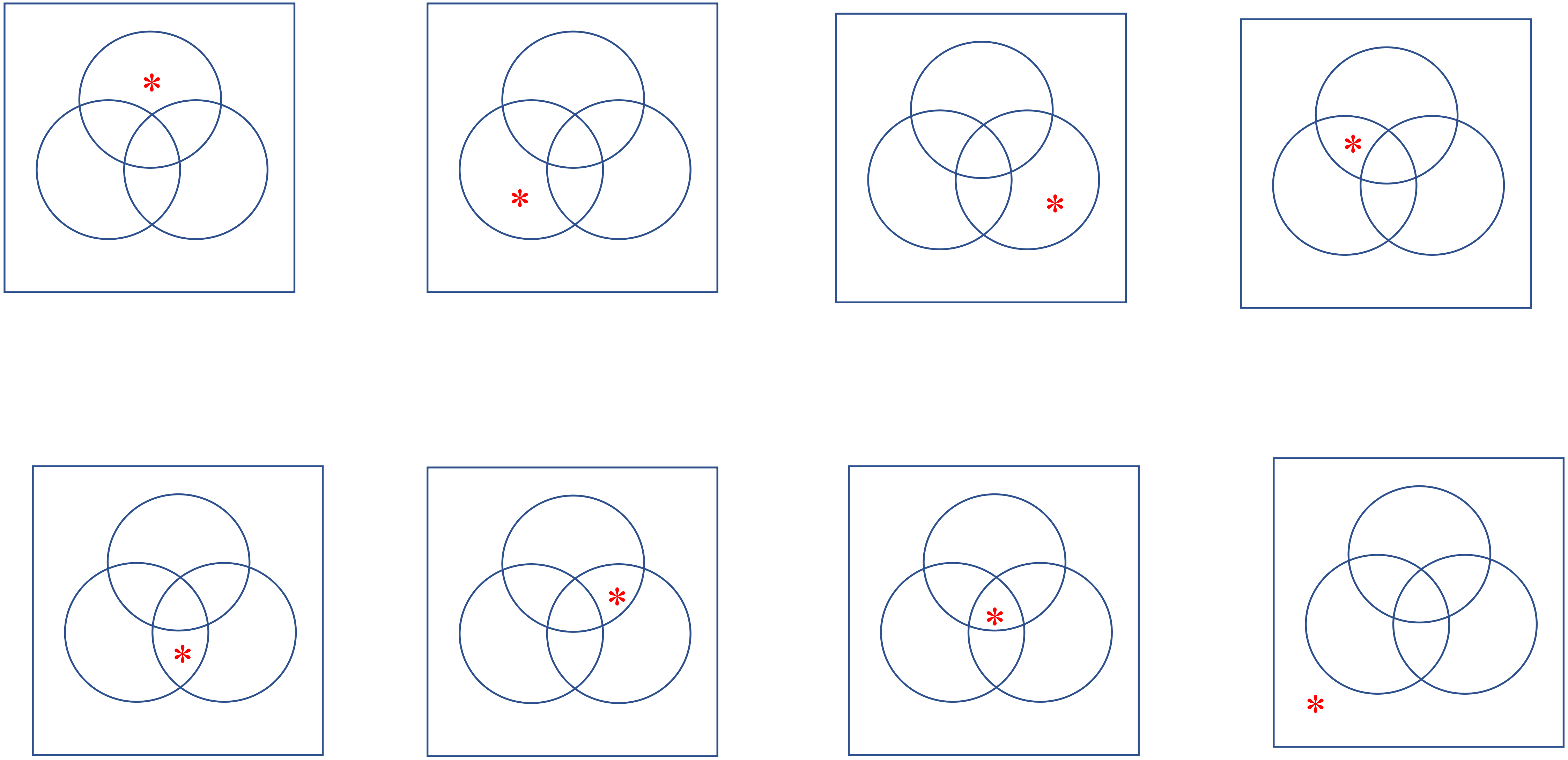
Suppose we are checking BVA(every boy, his).

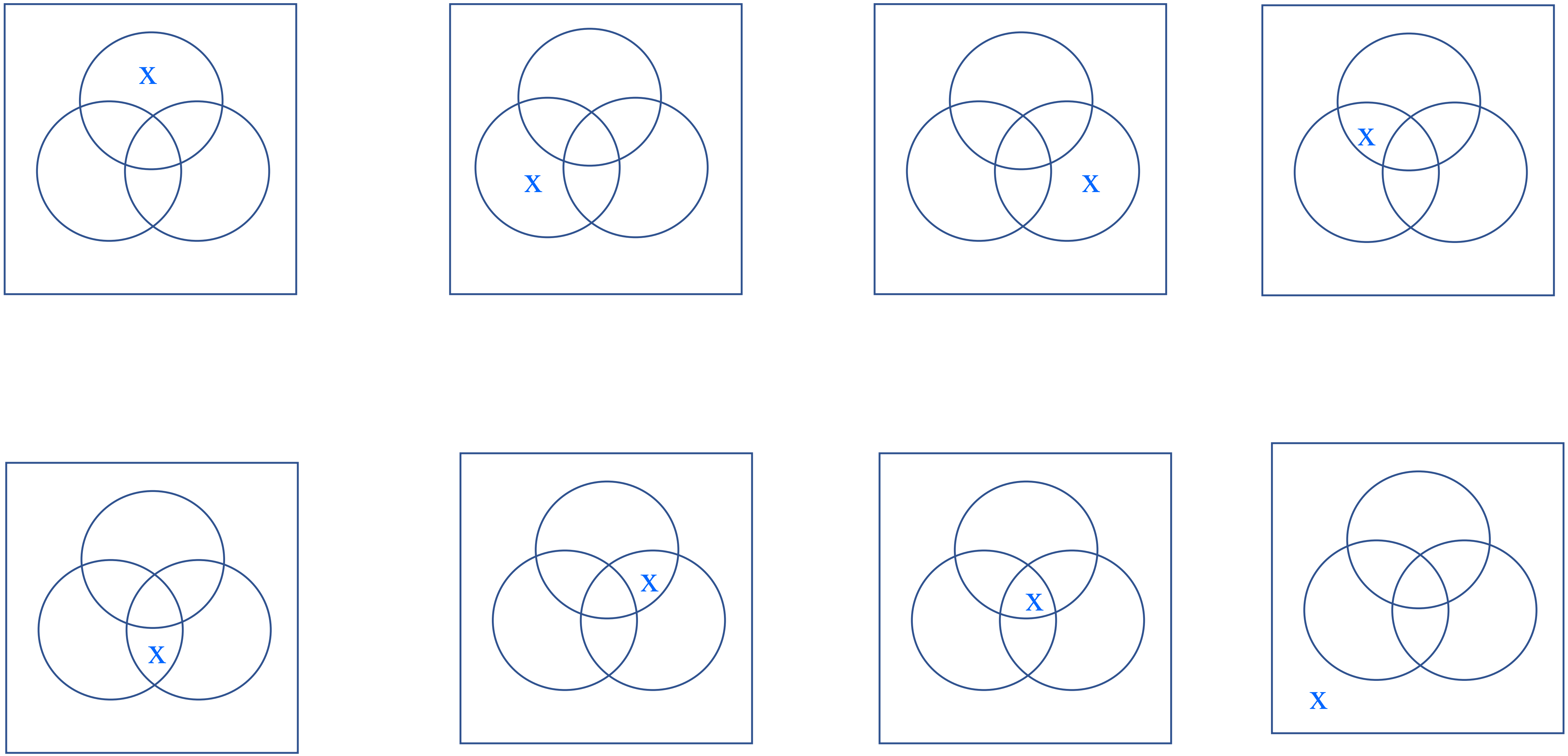
Speaker 1



What places a speaker in a particular place in the Venn Diagram?
What determines the color of the speaker?

What places Speaker 1 and Speaker 2 in their respective places, for example?




x
x
x
x
x
x
x
x

Being in the middle intersection means:
(i) the speaker is attentive and there are no "compatibility issues" (between the speaker and the experimental design) (based on the Subs test),
(ii) the X is [-XQrk] (based on the DR test), and
(iii) the Y is [-YQrk] (based on the Coref test),
for that speaker, when they judge the sentences.

Schema B (e.g., *His father praised every boy*) must correspond to a merge-generated set {his father, {praised, every boy}}. So, we have ¬CC(every boy, his).

$\neg$CC(every boy, his) $\wedge$ $\neg$XQrk $\wedge$ $\neg$YQrk $\rightarrow$ $\neg$BVA(every boy, his)

So, we make the prediction that if a speaker is in the middle intersection, they are not Red, i.e., they never say "Yes" on Schema B.

If they are Green, that is nice, but we do not predict those in the middle intersection are necessarily Green.

We have "... $\rightarrow \neg$BVA(X, Y)", but we don't have " ... $\rightarrow$ BVA(X, Y)".
We predict no Reds in the middle, but we do not predict Green in the middle.

Greens in the middle are instances of c-command detection (no longer just potential detection).

As noted, the Venn Diagram that we have seen should be understood as a composite of many Venn Diagrams, each of which looks like one of the 24 Venn Diagrams given in the preceding slides.

It is important to remember that the language faculty is internal to each of us and *definite* predictions in LFS is about an individual.

Our prediction of ¬BVA(X, Y) from ¬CC(X, Y) ∧ ¬XQrk ∧ YQrk is based on (1), the generalized revised hypothesis, plus (2), the hypothesis about correspondences between SVO and OSV to a merge-generated set.

(1) To have MR(X, Y), we must have: (i) X c-commands Y, (ii) X is [+XQrk], *or* (iii) Y is [+YQrk], with the '*or*' being inclusive '*or*'.

(2) SVO (Subject Verb Object) *must*, and OSV (Object Subject Verb) *can*, correspond to a merge-generated set {S, {V, O}}, where S c-commands O (and O does not c-command S).

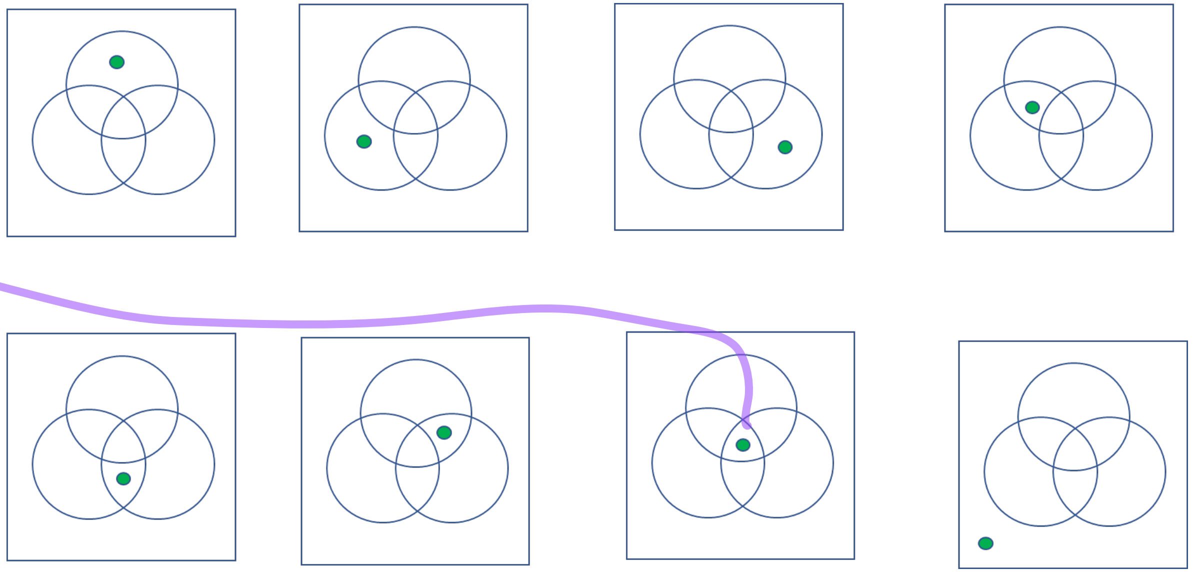
Both hypotheses, which our prediction is based on, make crucial reference to "c-command".

Our prediction is thus closely tied to the detection of c-command effects.

"Potential detection of c-command effects" = "C-command pattern"

"C-command pattern" = "No" on Schema B and "Yes" on Schema A (in a Main-Experiment)

"Detection of c-command effects" = "C-command pattern" by a speaker who is in the middle intersection of our Venn Diagram



"Detection of c-command effects" = "C-command pattern" as predicted

More accurately, what is predicted is the "No" judgment, not the "Yes" judgment.

The "No" judgment on Schema B (where X does not c-command Y in the merge-generate set, by hypothesis) is predicted for a speaker who is in the middle intersection of our Venn Diagram, since being in the middle intersection "means" that X is [-XQrk] and Y is [-YQrk] for that speaker and the speaker is attentive when they are making the relevant judgments.

The “Yes” judgment on Schema A (which, by hypothesis, can correspond to a merge-generated set where X c-commands Y), though not predicted, is an integral part of c-command detection. Because the “No” judgment on Schema B alone can be due to something other than the absence of the requisite c-command. (BVA is not possible in ‘John’s father praised every boy’, but that is not due to ‘every boy’ not c-commanding ‘John’. Even with the possibility of ‘every boy’ c-commanding ‘John’, as in ‘John’s father, every boy praised, the BVA is not possible, unlike ‘His father, every boy praised’.)

It is the “Yes” judgments on Schema A that differs minimally from Schema B—only in that X c-commands Y in Schema A but not in Schema A—that leads us to conclude that the “No” judgment on Schema B is indeed due to X not c-commanding Y. Based on “$\neg$CC(X, Y) $\wedge$ $\neg$XQrk $\wedge$ $\neg$YQrk $\rightarrow$ $\neg$MR(X, Y)” (which is the contrapositive of our hypothesis “MR(X, Y) $\rightarrow$ CC(X, Y) $\vee$ XQrk $\vee$ YQrk”), we take the “No” judgment on Schema B as indicating, though it does not entail, that the X is [-XQrk] and the Y is [-YQrk]. The “Yes” judgments on Schema A cannot be due to XQrk or YQrk; it must then be due to CC(X, Y).

The “Yes” judgment on Schema A, accompanied by the “No” judgment on the corresponding Schema B, is considered the detection of c-command effects for this reason.

In summary, we do not predict the “Yes” judgment on Schema A. But it is, ultimately, the “Yes” judgment on Schema A that gives us c-command detection, of course, accompanied by the “No” judgment on the corresponding Schema B, which we do predict.

Successful detection of c-command effects will provide support for the Merge hypothesis because "c-command" is defined in terms of Merge (a c-commands b: a is merged with something that contains b).

The hypothesis about Merge is for the initial state of the language faculty, hence about a *universal* aspect of the human mind.

Successful detection of c-command effects by the basic scientific method ("Guess-Compute-Compare"), therefore suggests that part of the mind can be studied by the basic scientific method!!

“Factual knowledge” vs. “Comprehension”

We have emphasized that language faculty science focuses on *individuals*.

Our prediction is about an *individual*, not about a group of speakers.

Let us consider, very briefly, what kind of a picture we might see if we considered "groups" instead of individuals.

The result of BVA[31]-37 (BVA(every N, his)) over the years, before the fall of 2022, focusing on native speakers of English.

No reference to the DR test, the Core test, or Subs

EPSA [31] : BVA in English
Experiment #37 -- Summary

The values for the Examples of the Schema groups and/or the Lexical groups that are "Excluded" are not shown.

participant list : **pers-r76.lst**

straight file : **s20.lst** | Yes-or-No (in sets) | Yes-or-No (one each) | --- | --- |

EPSA [31]-#37 < english > (Total 854 participants ; 7351 answers) ... as of Nov/15/2022

	Schema A	Schema B	Schema C
% of YES Answers	**43.5 %**	**17.8 %**	**89.6 %**
Number of Answers	2446	2450	2455

Based on this, one may "predict" that we will get a similar result, especially on Schema B, with a new group of speakers.

The result of BVA[31]-37 (BVA(every N, his)) in the fall of 2022, focusing on the native speakers of English.

No reference to the DR test, the Core test, or Subs

EPSA [31] : BVA in English
Experiment #37 -- Summary

The values for the Examples of the Schema groups and/or the Lexical groups that are "Excluded" are not shown.

participant list : **pers-r78.lst**

straight file : **s20.lst** | Yes-or-No (in sets) | Yes-or-No (one each) | --- | --- |

EPSA [31]-#37 < english > (Total 51 participants , 927 answers) ... as of Nov/15/2022			
	Schema A	**Schema B**	**Schema C**
% of YES Answers	**39.7 %**	**18 %**	**93.8 %**
Number of Answers	309	307	311

The result of BVA[31]-37 (BVA(every N, his)) over the years, prior to the fall of 2022, focusing on native speakers of English.

EPSA [31] : BVA in English
Experiment #37 -- Summary

The values for the Examples of the Schema groups and/or the Lexical groups that are "Excluded" are not shown.

participant list : **pers-r76.lst**

straight file : **s20.lst** | Yes-or-No (in sets) | Yes-or-No (one each) | --- | --- |

EPSA [31]-#37 < english > (Total 854 participants ; 7351 answers) ... as of Nov/15/2022

	Schema A	Schema B	Schema C
% of YES Answers	43.5 %	17.8 %	89.6 %
Number of Answers	2446	2450	2455

The date shown (Nov/15/2022) is the date when these charts were created. The top chart excludes the 51 participants in the fall of 2022, and the bottom chart is only about those participants in the fall of 2022.

Very close!!

The result of BVA[31]-37 (BVA(every N, his)) in the fall of 22f, focusing on the native speakers of English.

EPSA [31] VA in English
Experiment 7 -- Summary

The values for the Examples of the Schema groups and/or the Lexical groups that are "Excluded" are not shown.

participant list : **pers-r78.lst**

straight file : **s20.lst** | Yes-or-No (in sets) | Yes-or-No (one each) | --- | --- |

EPSA [31]-#37 < english > (Total 51 participants ; 927 answers) ... as of Nov/15/2022

	Schema A	Schema B	Schema C
% of YES Answers	39.7 %	18 %	93.8 %
Number of Answers	309	307	311

The results shown in the preceding slide may be impressive to those interested in "factual knowledge" but not to those interested in "comprehension".

Based on our hypotheses, we deduce "0" on Schema B (as the %Value(STB)); there should not be any individual in the center intersection who accepts Schema B sentences of BVA-Main to any extent, provided that they pass the Subs test, and are Green with the DR test and the Coref test.

That is how "definite" our prediction is. In language faculty science, Guess-Compute-Compare is applied to an *individual*, not to a group of speakers.

We don't have hypotheses about a group of speakers, so we don't make predictions about a group.

We certainly do not deduce the %Value(STB) being "18%" about a group of speakers from any hypotheses.

The difference between the group-based “approach” and the individual-based “approach” is an instance of “factual knowledge” vs. “comprehension”, as addressed in Einstein’s remarks in his Foreword to *Dialogue Concerning the Two Chief World Systems*.

[Galileo's] endeavors are not so much directed at "factual knowledge" as at "comprehension."

"Comprehension" is like understanding things by the Guess-Compute-Compare method.

From Einstein’s Foreword to *Dialogue Concerning the Two Chief World Systems—Ptolemaic & Copernican* by Galileo Galilei

“It has often been maintained that Galileo became the father of modern science by replacing the speculative, deductive method with the empirical and experimental method.

I believe, however, that this interpretation would not stand close scrutiny.

There is no empirical method without speculative concepts and systems; and there is no speculative thinking whose concepts do not reveal, on closer investigation, the empirical material from which they stem.

To put into sharp contrast the empirical and the deductive attitude is misleading, and was entirely foreign to Galileo.

Actually it was not until the nineteenth century that logical (mathematical) systems whose structures were completely independent of any empirical content had been cleanly extracted.

Moreover, the experimental methods at Galileo's disposal were so imperfect that only the boldest speculation could possibly bridge the gaps between the empirical data. (For example, there existed no means to measure times shorter than a second.) [Continuing to the next slide.]

The antithesis Empiricism vs. Rationalism does not appear as a controversial point in Galileo's work.

Galileo opposes the deductive methods of Aristotle and his adherents only when he considers their premises arbitrary or untenable, and he does not rebuke his opponents for the mere fact of using deductive methods.

In the first dialogue, he emphasizes in several passages that according to Aristotle, too, even the most plausible deduction must be put aside if it is incompatible with empirical findings.

His endeavors are not so much directed at "factual knowledge" as at "comprehension".

But to comprehend is essentially to draw conclusions from an already accepted logical system." (pp. xvii-xviii)

(This is how Einstein ends his Foreword.)

A Side Note:

AI research seeks “factual knowledge” (collecting tons of “data” and make “predictions” about new data, without deduction from “first principle”) while LFS seeks “comprehension”, by the “Guess-Compute-Compare” method.

Continuation of Slide 7

The LFStist also examines new MRs (not just BVA, DR, and Coref), and checks a number of new predictions under new hypotheses about those new MRs (and related issues), further expanding their empirical coverage.

What we have discussed is the very basic part of LFS research. Basic, but fundamental, because we are checking c-command effects, assumed to be observable in any speaker of any language, according to the Merge hypothesis.

Starting with the very basic schemata (sentence patterns).

Checking is done by the "Guess-Compute-Compare" method (=the basic scientific method)!

The significance of success of LFS

The main *general* thesis of the language faculty science research is that it is possible to accumulate knowledge about part of the mind (the language faculty) by the basic scientific method (by the "Guess-Compute-Compare method) (so in a way very much as in physics).

This is a remarkable result if true; very few people in the world know about this!

The three lectures on language faculty science are meant to provide a brief introduction to LFS.

More detailed, but less up-to-date, exposition is found in *The Theory and Practice of Language Faculty Science* (10.1515/9783110724790).

Revised versions of Chapters 4, 5, and 6 therein are available at:
Chapter 4: The key tenets of language faculty science
Chapter 5: Detection of c-command effects
Chapter 6: Replication: Predicted Correlations of Judgments in Japanese

Sections 1 and 2 of Chapter 9: Language faculty science and physics also provides introductory remarks about LFS.

The logic of LFS is not fully presented in these lectures.
How the basic idea of LFS is applied to various "empirical domains" and how further conceptual and theoretical articulation is being sought will have to be addressed in other lectures.