Language Learning & Technology

Guidelines on Conducting and Reporting on CALL Research

The following slides are based on workshops presented at CALICO Conferences by (current and previous) editors of LLT.

Dorothy Chun
Trude Heift
Carla Meskill
Irene Thompson

© 2017, Language Learning & Technology

© Language Learning & Technology

Outline

Part 1: Statement of the problem

Part 2: Theoretical perspective

Part 3: Research methodology

Part 4: Data analysis and interpretation of results

Part 5: Preparing to publish & specific concerns

© Language Learning & Technology

Part 1

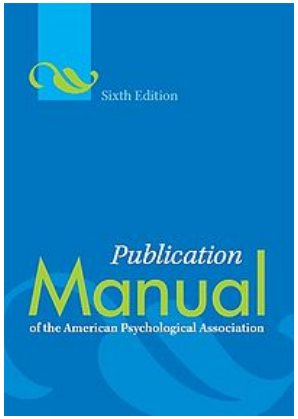
Statement of the problem

© Language Learning & Technology

Preparing to do research

Designing and reporting research

Chapter 1, Publication Manual of the American Psychological Association (APA), 6th ed. (2009), **2nd printing**.



Journal articles

- Empirical studies
- Lit review articles
- Theoretical articles
- Methodological articles
- Case studies

© Language Learning & Technology

Empirical studies

- Main components
  - Introduction
  - Literature review
  - Method
  - Results
  - Discussion

- Quantitative vs. qualitative studies
  - e.g., see LLT Guidelines

© Language Learning & Technology

LLT guidelines

An **Introduction** that

- states the problem to be investigated
- contextualizes the research by describing the underlying theoretical framework and reviewing previous studies
- defines the variables and research hypotheses

A **Method Section** that describes

- the participants (e.g., demographics, selection criteria, and group assignment)
- the materials (e.g., task[s], equipment, instruments, including a discussion of their validity and reliability, if appropriate)
- the procedures employed in the study such as treatment(s)

A **Results Section** that includes

- graphs and tables that help to present and explain the results
- descriptive and inferential statistics used to analyze the data, including the following:
    - name of the statistic used and in the case of an uncommon statistical procedure, a reference to a discussion of the procedure
    - statistical significance of the results obtained
    - measures of effect sizes
    - how all necessary assumptions were met

A **Discussion Section** that includes

- an interpretation of the results
- an explanation of the results, including alternative explanations when appropriate
- a statement relating the results obtained in the study to original hypotheses
- theoretical implications
- limitations of the study

A **Conclusion** that includes

- general implications of the study
- limitations of the study
- suggestions for further research

References

Appendices

© Language Learning & Technology

Other types of studies

- Literature review articles
  - Include meta-analyses
  - Critical evaluations of published materials
- Theoretical articles
  - Ordinarily present a new theory
  - May critically analyze existing theories
- Methodological articles
  - New approaches to data analysis
- Case studies
  - Often longitudinal, small “n,” naturalistic observation

© Language Learning & Technology

Ethics of research

- IRB (Institutional Review Board) applications
  - In Office of Research at UCSB, now online
  - Check online for forms at your university
- Human Subjects Informed Consent
  - Application must be filed and approved prior to any data collection
  - Researchers are often expected to complete a training or tutorial (nowadays often online)

© Language Learning & Technology

Rules of thumb for IRB

- State that the goal is to investigate and improve instruction.
- If applicable, state that the study falls within normal course curriculum; content of course is not altered.
- Emphasize that there is no physical or psychological risk to study participants.
- List who will have access to the data.
- State that students will never be identified by name in research reports.
- Explain how students will benefit from the study.

© Language Learning & Technology

Copyright issues

- For CALL research and development, consult the Fair Use Guidelines for Educational Multimedia.
- To use copyrighted materials in your research, consult your university's library (e.g., see UCSB's Library).
- To incorporate information taken from the Internet, link directly to URLs rather than copying **or** request permission from copyright holder.

© Language Learning & Technology

Evaluation of CALL research

- Joint Policy Statements of CALICO ("Scholarly activities in CALL: Development, pedagogical innovations and research"), EuroCALL ("Research Policy Statement"), and IALLT (1999).

- MLA Guidelines for Evaluating Work with Digital Media in the Modern Languages (2012).

© Language Learning & Technology

Main CALL Journals

- ***CALICO***
  Bryan Smith, Mat Schulze

- ***CALL***
  Jozef Colpaert

- ***ReCALL***
  Françoise Blin, Alex Boulton

- ***System***
  Xuesong Gao, Marta González-Lloret, Ursula Stickler, Lawrence Jun Zhang

LANGUAGE LEARNING & TECHNOLOGY
a refereed journal for second and foreign language educators

© Language Learning & Technology

References

- Bitchener, J. (2009). *Writing an applied linguistics thesis or dissertation: A guide to presenting empirical research*. London: Palgrave Macmillan.
- Arnold, N., & Ducate, L. (Eds.). (2011). *Present and future promises of CALL: From theory and research to new directions in language teaching.* San Marcos, TX: Computer Assisted Language Instruction Consortium.
- Guikema, J. P., & Williams, L. F. (Eds.). (2014). *Digital Literacies in Foreign and Second Language Education*. San Marcos, TX: Computer Assisted Language Instruction Consortium (CALICO).
- Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer-assisted language learning*. Mahwah, NJ: Erlbaum.
- Mackey, A., & Gass, S. M. (2016). *Second language research: Methodology and design,* 2nd ed. New York: Routledge.
- Norris, J., & Ortega, L. (2003). Defining and measuring SLA. In C. J. Doughty, & M. H. Long, (Eds.), *The handbook of second language acquisition* (pp. 717-761). Malden, MA: Blackwell.
- Smith, B., & Lafford, B. (2009). The evaluation of scholarly activity in computer-assisted language learning. *The Modern Language Journal, 93* (Focus issue), 868-883.
- Tarone, E. E., Gass, S. M., & Cohen, A. D. (Eds.). (1994). *Research methodology in second-language acquisition*. Mahwah, NY: Erlbaum.

© Language Learning & Technology

Part 2

Theoretical perspective

© Language Learning & Technology

TAOIST WRITER, LIEH-TSE

A man noticed that his ax was missing.
Then he saw his neighbor's son pass by.
The boy looked like a thief and behaved like a thief.
Later that day, the man found his ax where he had left it the day before.
The next time he saw his neighbor's son, the boy looked, walked and behaved like an honest, ordinary boy.

© Language Learning & Technology

© Language Learning & Technology

Perspectives

How we see and interpret the world depends on our position, our perspective.

How we talk and write about the world is shaped by our position, our perspective.

© Language Learning & Technology

Framing

Lakoff's *Don't Think of an Elephant*

- Metaphors -
Cognitive structures
that guide our perceptions

© Language Learning & Technology

Frames

When undertaking and reporting on primary,

empirical research:

a clear theoretical position and

articulation of assumptions

© Language Learning & Technology

Perspectives

Traditionally single journal, single POV

When you open some journals, perspectives and assumptions are assumed under title/ editorship

© Language Learning & Technology

Perspectives

Nowadays, there is a wide range of perspectives from which to tackle pedagogical phenomena

The range of perspectives for inquiry broaden greatly

© Language Learning & Technology

Assumptions about mind, language, learning

Learning is a matter of cause and effect **FRAME**

Social, cultural, historical and political (critical) forces shape the learning **FRAME**

© Language Learning & Technology

Theoretical frame

= a conceptual tool that can move an inquiry forward toward deeper levels of understanding.

© Language Learning & Technology

Theoretical frame - cont.

- Determines the research problem
- Shapes and defines the scope and direction of the literature review
- Words the research questions
- Steers methods
- Guides analysis
- FRAMES INTERPRETATION
- FRAMES conclusions
- SIGNALS THE RELEVANCE TO THE FIELD OR AREA OF INQUIRY

© Language Learning & Technology

Resources

Anfara, V. and Mertz, N. (2006). *Theoretical frameworks in qualitative research. Thousand Oaks*, CA: Sage.

Boote, D. and Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, *34*(6), 3-15.

Creswell, J. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. New York: Sage.

Egbert, J. and Petrie, G. (2005). *CALL research perspectives.* Mahwah, NJ: Lawrence Erlbaum.

Harre, R. and van Langenhove, L. (1999) Varieties of positioning. *Journal of the Theory of School behavior*, *21*, 393-408.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*(3), 465-491.

Lakoff, G. (2004). *Don't think of an elephant*. White River Junction, VT: Chelsea Green Publishing.

Levy, M. (2015). The role of qualitative approaches to research in CALL contexts: Closing in on the learner's experience. *Calico Journal, 32*(3), 554-568.

Lewis, C., Enciso, P., and Moje, E. (2007). *Reframing socicultural research on literacy: Identity, agency, and power*. New York: Routledge.

Nunan, D., & Bailey, K. M. (Eds.). (2009). *Exploring Second Language Classroom Research: A Comprehensive Guide*. Boston: Heinle Cengage Learning.

© Language Learning & Technology

Sample A

Title: Achievement and Retention of Spanish Presented Via Videodisc in Linear, Segmented and Interactive Modes

Abstract: This study investigated the effects on achievement and retention of a beginning Spanish instructional videodisc entitled Zarabanda, presented over a two week period in linear, segmented, and interactive modes. Students (N = 92) enrolled in a beginning Spanish course at the U.S. Air Force Academy were randomly Assigned to three treatment groups and a control group. Group 1 (n = 23), the linear videodisc instruction group (descriptive), watched the Zarabanda videodisc in a passive manner. Group 2 (n = 23), the segmented videodisc instruction group (experimental), was presented the same material as Group 1 with the addition of inserted true/false and multiple choice questions at selected breakpoints in the story line. Group 3 (n = 23), the interactive videodisc instruction group (experimental), was presented materials in an interactive mode. The lesson was interrupted by the same questions at the same breakpoints as Group 2 with the additional benefit of feedback on incorrect choices, vocabulary lists, video replay and hint options for remediation, and explanatory statements on correct choices.

© Language Learning & Technology

Sample B

Title: Using Native Speakers in Chat

Abstract: SLA research indicates that negotiation promotes interlanguage development and that learners are most likely to negotiate if opportunities for oral interaction are provided. In the case of campus-based students, learners' progress is supported and monitored mainly through classroom interactions. If students do not attend classes on campus, how do they gain the reported benefits of oral interaction? Recent studies indicate that chatting provides opportunities for the negotiation of meaning, as occurs in oral interaction. However, most of these have been conducted on interactions between learners, with teacher supervision, often in task-based instructional settings. This study considers implications for distance language learning of negotiations by a group of intermediate learners of Italian interacting in dyads on a Web based Italian native speaker (NS) chat program. The research specifically explores (a) whether live chat with native speakers offers opportunities for negotiation of meaning in open ended tasks carried out in single session interactions with unfamiliar NS without teacher supervision, (b) the principal triggers for negotiation and modification of interlanguage in these interactions, and (c) whether public NS chat rooms are likely to offer an optimal environment for SLA, even for learners studying at a distance who need to chat without supervision. Chat logs indicate that learners do in fact negotiate for meaning and modify their interlanguage when engaged in open ended conversational tasks with unfamiliar interlocutors, with lexical and structural difficulties triggering most negotiations. Though further research needs to probe whether these negotiations and modifications lead to acquisition in the longer term, they would be particularly valuable for distance learners who need opportunities to negotiate within authentic target language contexts.

© Language Learning & Technology

Sample C

Title: Computer Assisted Second Language Vocabulary Acquisition

Abstract: During the initial stages of instructed L2 acquisition students learn a couple thousand, mainly high frequency words. Functional language proficiency, however, requires mastery of a considerably larger number of words. It is therefore necessary at the intermediate and advanced stages of language acquisition to learn a large vocabulary in a short period of time. There is not enough time to copy the natural (largely incidental) L1 word acquisition process. Incidental acquisition of the words is only possible up to a point, because, on account of their low frequency, they do not occur often enough in the L2 learning material. Acquisition of new words from authentic L2 reading texts by means of strategies such as contextual deduction is also not a solution for a number of reasons. There appears to be no alternative to intentional learning of a great many new words in a relatively short period of time. The words to be learned may be presented in isolation or in context. Presentation in bilingual word lists seems an attractive shortcut because it takes less time than contextual presentation and yields excellent short term results. Long term retention, however, is often disappointing so contextual presentation seems advisable. Any suggestions how to implement this in pedagogic contexts should be based on a systematic analysis of the two most important aspects of the L2 word learning problem, that is to say, selecting the relevant vocabulary (which and how many words) and creating optimal conditions for the acquisition process. This article sets out to describe a computer assisted word acquisition programme (CAVOCA) which tries to do precisely this: the programme operationalises current theoretical thinking about word acquisition, and its contents are based on a systematic inventory of the vocabulary relevant for the target group. To establish its efficiency, the programme was contrasted in a number of experimental settings with a paired associates method of learning new words. The experimental results suggest that an approach combining the two methods is most advisable.

© Language Learning & Technology

Sample D

Title: Triadic Scaffolds: Tools for Teaching English Language Learners withComputers

ABSTRACT: Active communication with others is key to human learning. This straightforward premise currently undergirds much theory and research in student learning in general, and in second language and literacy learning in particular. Both of these academic areas have long acknowledged communication's central role in successful learning with the exact intricacies of instructional conversations and the forms these take having been the focus of close analysis (Cazden, 1988; Gee, 2001; Nystrand, Gamoran, Kachur, & Prendergast, 1997; Tharp & Galimore, 1991; van Lier, 2000). In this examination of computer-supported classroom discourse, specific forms of instructional conversation employed by a veteran elementary teacher of beginning-level English language learners (ELLs) are examined. The focal teacher orchestrates instructional conversations around computers with children whose immediate needs are to learn the English language, specifically the "language of school" and the concomitant social complexities implied in order to participate in mainstream instructional activity. With these goals shaping language and literacy activity, their ESOL (English for speakers of other languages) teacher makes use of the computer to capture, motivate, and anchor learner attention to, and render comprehensible the target language they hear and see on and around the computer screen. The anatomy of the activity she orchestrates around the computer and the language she uses to support it -- labeled here as *triadic scaffolds* -- are the focus of analysis. Forms and functions of triadic discourse (teacher, learner, computer) are examined for their potential unique role in second language and literacy instruction.

© Language Learning & Technology

Part 3

Research methodology

© Language Learning & Technology

Steps in conducting CALL Research

1. Choose the theory/frame
2. Formulate the research hypothesis/hypotheses based on theory
3. Develop the study design
4. Develop the sampling plan
5. Define the variables and control for extraneous ones
6. Choose an appropriate measurement instrument by keeping its validity and reliability in mind
7. Collect and analyze data

© Language Learning & Technology

Step 1: Choose theory

Examples:

- Dual-coding theory (Paivio)

Visual and verbal information are processed in two different channels.

- Cognitive theory of multimedia learning (Mayer)

Auditory and visual channels have limited capacity for processing information. Learning is an active process of filtering, selecting, organizing, and integrating information.

- Social constructivist theory (Vygotsky)

Culture and social context are critically important for learning to take place.

© Language Learning & Technology

Step 2: State null hypothesis (H_o)

- Base H_o on your theoretical framework
- Your study should be designed so as to support or refute H_o
- Examples of null hypotheses:

  "Visual representations do not support learning L2 nouns and verbs"

  "Email exchanges with native speakers do not influence the development of sociolinguistic norms in the L2"

© Language Learning & Technology

Step 3: Choose study design

- Experimental

Most rigorous of all designs and the best method for coming to conclusions about cause and effect.

- Quasi-experimental

Lacks the control of true experimental design and is more susceptible to alternative explanation of findings.

- Qualitative

Examination of naturally occurring phenomenon that accounts for contextual complexities.

© Language Learning & Technology

Experimental (two-group) design

- "Gold standard" since it has the strongest internal validity.
- It uses two groups that are as equivalent to each other as possible. Random assignment is key.
- Experimental (treatment) group gets treatment, control (comparison) group does not. Otherwise, the groups are treated the same.
- Difficult to carry out in real-world instructional settings.

© Language Learning & Technology

Quasi-experimental design

- Looks like an experimental design but lacks random assignment.
- Poses threat to internal validity, i.e., we can erroneously conclude that our treatment had an effect when it didn't, or that it had no effect when it did.

© Language Learning & Technology

Common Quasi-Experimental Designs

- Nonequivalent groups

• uses intact groups (e.g., two classrooms) as the treatment and comparison groups. Subject to internal validity threat since groups may be different *prior* to treatment.

- One-Group Posttest-Only

• lacks a pretest baseline or a comparison group, making it impossible to come to valid conclusions about the treatment effect. Results may be due to any number of reasons.

- One-Group Pretest-Posttest

is subject to internal validity threats, such as

▪ history (events between pre- and posttest)
▪ maturation (students would have learned anyway)
▪ regression toward the mean (the tendency of extremes to revert toward the average)
▪ testing (the learning effect of pretest on posttest)

© Language Learning & Technology

Step 4: Develop a sampling plan

Research aims at being able to generalize findings to more than just the study participants.

- Random selection

Refers to selection of the sample. It is most related to *external validity* (generalizability) of results since it assumes that the selection accurately represents the population from which it was drawn. SLA researchers often use their own classes.

- Random assignment

Refers to assignment of subjects to the groups. It is most closely related to *internal validity* since it ensures that the groups are equivalent *prior* to treatment. SLA researchers often use intact classes.

© Language Learning & Technology

Variables

Independent variable, or factor, is one that is being studied and whose effects are being manipulated, e.g., type or length of treatment.

Dependent variable, or measure, is the one that is presumed to be affected by the independent variable and that is used to measure its effect, e.g., number of words recalled as a result of exposure to types of different cues.

Extraneous variable is one that is not accounted for in the design, e.g.,

- proficiency level, L1, age, and gender of subjects
- instructional setting
- time on task
- instructor
- instructional materials

© Language Learning & Technology

Step 5: Define and control Variables

- Extraneous Variables
  - proficiency level, age, gender, setting
  - time on task
  - instructor
  - instructional materials
- Use Factorial designs

that include extraneous variables in the research design, e.g., gender. This eliminates it as a potential uncontrolled variable.

© Language Learning & Technology

Examples of Statistical Controls

- **Analysis of Covariance (ANCOVA)** statistically removes the effect of the pretest so that you can just look at the difference in the posttest measurements as reflecting only the effects of the treatment.

- **Partial regression** makes it statistically possible to look at the relationship between two variables, e.g., between grammar and reading comprehension, while controlling for the differences related to scores on another variable, e.g., vocabulary.

© Language Learning & Technology

Step 6: Establish validity and reliability

The best instrument is one that is both valid and reliable.

Validity and reliability do not always go together.

- An instrument may be reliable but not valid because it may be measuring the wrong construct consistently.
- An instrument may be valid but not consistent in measuring the construct.

© Language Learning & Technology

Types of validity

- Construct validity

degree to which inferences can legitimately be drawn from the results of the study to the theoretical constructs on which the study is based.

- External validity

degree to which the results of the study can be generalized to other contexts (people, places, times).

- Internal validity

degree to which observed results can be attributed to the treatment and not to other alternative explanations. It is not relevant in most observational or descriptive studies but is of primary consideration in studies that assess effects of treatments.

Validity types

© Language Learning & Technology

Types of construct validity

- **Face validity**
  Instrument is a good representation of the construct "on its face".
- **Content validity**
  Instrument adequately covers content of the treatment.
- **Predictive validity**
  Instrument predicts what it theoretically should be able to, e.g., oral interview should be able to predict a person's performance in comparable F2F communicative situations.
- **Concurrent validity**
  Instrument distinguishes between groups that it should theoretically be able to distinguish, e.g., technology attitude questionnaire should distinguish between computer-savvy and computer-illiterate subjects (whose background has been established independently).
- **Convergent validity**
  Degree to which the measurement is similar to other similar measurements, e.g., your homemade test produces results similar to those on a standardized test whose validity is known.

© Language Learning & Technology

Reliability of Measurement Instrument

Any measurement score = true ability + random error. We want to minimize effects of random error.

Reliability must be reported.

- **Inter-rater reliability**

Extent to which two or more coders or raters agree. It addresses the consistency of the implementation of a rating system.

- **Test-retest reliability**

The same test is administered to the same sample on two different occasions. The correlation between the two sets of scores is the estimate of reliability. More suitable for surveys than for educational tests.

© Language Learning & Technology

Reliability of Measurement Instrument - cont.

- **Parallel-forms reliability**

- Consistency of the results from two tests designed to measure the same content.
  - develop a large set of questions that address the same construct
  - divide the questions into two sets
  - administer one set as pretest and the other as posttest
  - compute correlation between the two sets of scores to estimate reliability

- **Internal consistency reliability**

- We administer a single test and look at the consistency of results across all items in a test. There are many internal consistency measures:
  - average inter-item correlation
  - average item-to-total correlation
  - split-half correlation
  - odd-even correlation
  - *Cronbach's alpha* (for continuous measures), *Kuder-Richardson* (KR-20) for dichotomous data

© Language Learning & Technology

Step 7: Collect and analyze data

- Involve a statistician in the design of your study to make sure that the data you are planning to collect can be analyzed statistically.
- Report both descriptive and interpretive statistics.
- Descriptive statistics are relatively easy to report (you can do it), interpretive statistics are not (you probably cannot do it yourself).
- Take a course in statistics. This will not make you a statistics maven but it might help you understand what your statistician is saying.

© Language Learning & Technology

Conclusions

- No single study will ever be sufficient for understanding a particular phenomenon.

- Multiple studies asking the same research question are more likely to lead to the truth through replications and variations.

- It is very important to clearly describe and document your methodology so that your study can be replicated or varied.

© Language Learning & Technology

Part 4

Data analysis and interpretation of results

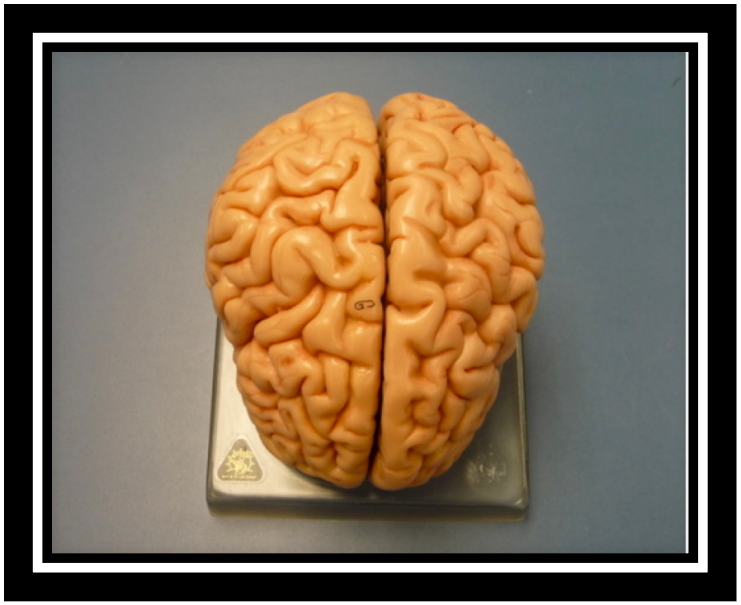
© Language Learning & Technology

Sample study: Research question

- Do pictorial cues facilitate the acquisition of L2 high- and low-imagery words?

© Language Learning & Technology

Imageability



high imageability (concrete):

das Gehirn – the brain

low imageability (abstract):

die Leistung – the performance

© Language Learning & Technology

Data

- 50 beginner learners of German who used Voka (Rimrott, 2009)
- Each study participant received 60 nouns (15 with a picture and 15 without for both low- and high-imagery nouns)
  - test items were taken from Berlin Affective Word List (Vö, Jacobs and Conrad, 2006)
    - words are rated for their imageability on a scale from 1 (low imageability) to 7 (high imageability) by 40 German speakers
      - Test items chosen: low-imagery: 1-2; high-imagery: 6-7
  - data collection
    - questionnaire
    - pre-test (a translation task): none of the students knew any of the test items prior to the study
    - study phase
      - 1. 22 sec: entire flashcard
      - 2. 16 sec: entire flashcard
      - 3. 4 sec: English word + 14 sec: entire flashcard
    - practice phase
      - students were asked to supply the correct word and they received feedback on their responses
    - immediate post-test (after practice phase) – identical to pre-test
    - delayed post-test (after one week) – identical to pre-test

© Language Learning & Technology

Test item with picture

Study Phase 3/3 | Word: 1/30

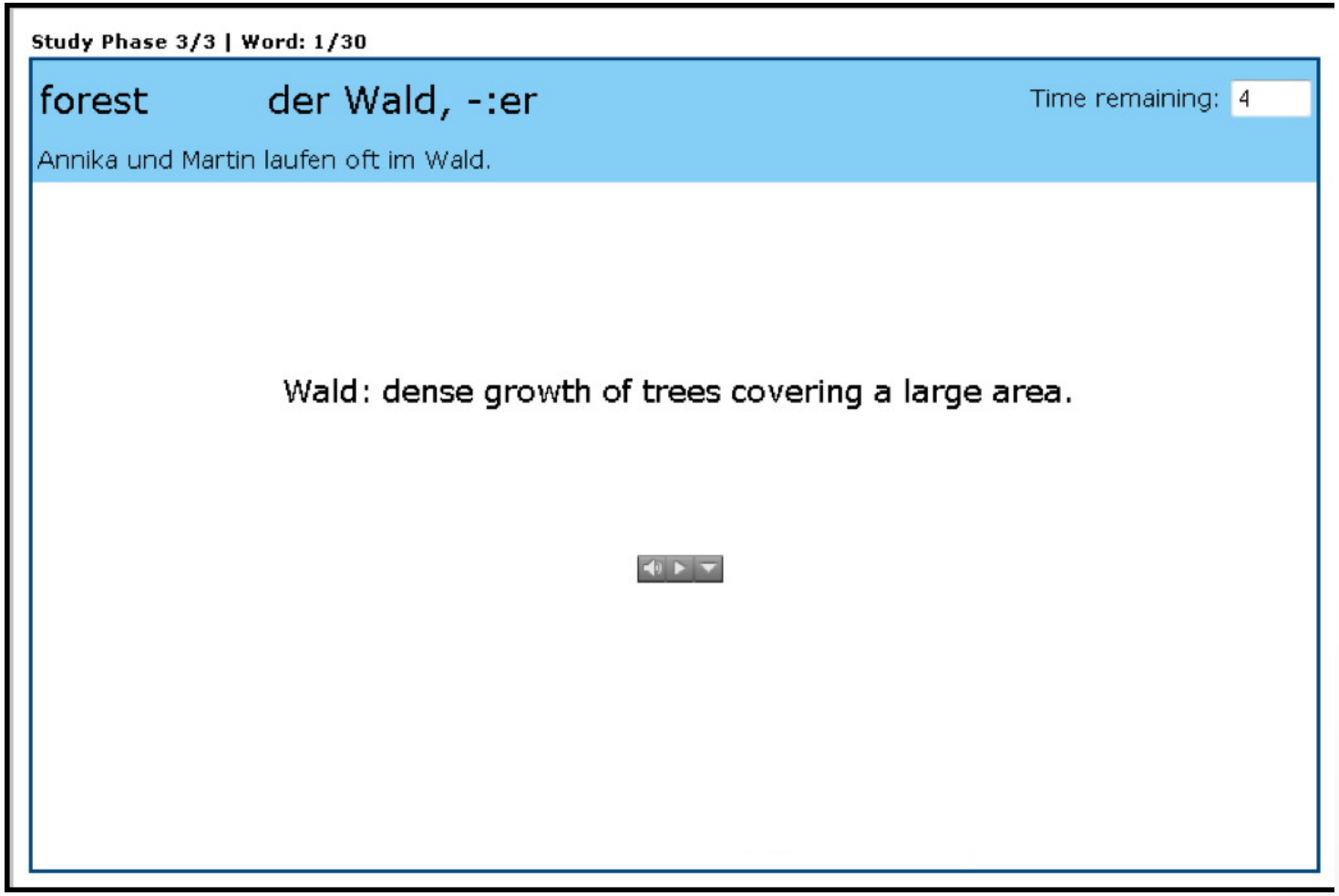
forest der Wald, -:er Time remaining: 4

Annika und Martin laufen oft im Wald.

Wald: dense growth of trees covering a large area.

© Language Learning & Technology

Test item without picture



© Language Learning & Technology

Data Analysis

- Given the data, how do we answer our research question?
  - Do pictorial cues facilitate the acquisition of L2 high- and low-imagery words?
- What additional research questions are prompted by the data?
  - Given the data, how can they be answered?

© Language Learning & Technology

Data Analysis

- Do pictorial cues facilitate the acquisition of L2 high- and low-imagery words?
  - high-imagery: picture – no picture
  - low-imagery: picture – no picture
- What additional research questions are prompted by the data?
  - What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?
  - Are high-imagery words ‘easier’ to acquire than low-imagery words?

© Language Learning & Technology

Statistics

- Working with a statistician…
- Do pictorial cues facilitate the acquisition of L2 high- and low-imagery words?
  - Descriptive statistics
    - are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures
  - Paired samples t-test
  - What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?
    - Repeated measures ANOVA
- Are high-imagery words ‘easier’ to acquire than low-imagery words?
  - Paired samples t-test

© Language Learning & Technology

Results: Paired Samples T-test

- Do pictorial cues facilitate the acquisition of L2 high- and low-imagery words?

			Mean	N	Std. Deviation	Std. Error Mean	t	Sig.
Post-test	Pair 1	Abstract_picture	9.80	50	5.067	.717	2.737	.009
		Abstract_no_picture	8.12	50	4.964	.702		
	Pair 2	Concrete_picture	12.30	50	3.290	.465	3.445	.001
		Concrete_no_picture	10.68	50	3.178	.449		
Delayed Post-test	Pair 1	Abstract_picture	5.60	50	5.395	.763	1.788	.080
		Abstract_no_picture	4.76	50	4.396	.622		
	Pair 2	Concrete_picture	8.60	50	5.107	.722	2.379	.021
		Concrete_no_picture	6.98	50	4.283	.606		

© Language Learning & Technology

Post-test results: repeated measures ANOVA

- What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?

Descriptive Statistics

	Mean	Std. Deviation	N
Abstract_picture	9.80	5.067	50
Abstract_no_picture	8.12	4.964	50
Concrete_picture	12.30	3.290	50
Concrete_no_picture	10.68	3.178	50

© Language Learning & Technology

Post-test results: repeated measures ANOVA

- What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?

Test of within-subject effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Groups	Sphericity Assumed	456.215	3	152.072	15.599	.000

© Language Learning & Technology

Post-test results: repeated measures ANOVA

- What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?

Pairwise Comparisons			
Groups	Groups	Std. Error	Sig.[a]
Abstract_picture	Abstract_no_picture	.614	.052
	Concrete_picture	.683	.004
	Concrete_no_picture	.622	.983
Abstract_no_picture	Abstract_picture	.614	.052
	Concrete_picture	.677	.000
	Concrete_no_picture	.655	.002
Concrete_picture	Abstract_picture	.683	.004
	Abstract_no_picture	.677	.000
	Concrete_no_picture	.470	.007
Concrete_no_picture	Abstract_picture	.622	.983
	Abstract_no_picture	.655	.002
	Concrete_picture	.470	.007

© Language Learning & Technology

Delayed post-test results: repeated measures ANOVA

- What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?

Descriptive Statistics

	Mean	Std. Deviation	N
Abstract_picture	5.60	5.395	50
Abstract_no_picture	4.76	4.396	50
Concrete_picture	8.60	5.107	50
Concrete_no_picture	6.98	4.283	50

© Language Learning & Technology

Delayed post-test results: repeated measures ANOVA

- What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?

Test of within-subject effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Groups	Sphericity Assumed	423.855	3	141.285	13.760	.000

© Language Learning & Technology

Delayed post-test results: repeated measures ANOVA

- What kind of help options (picture, no picture) are most effective for which word type (high-imagery, low-imagery)?

Pairwise Comparisons			
Groups	Groups	Std. Error	Sig.[a]
Abstract_picture	Abstract_no_picture	.470	.480
	Concrete_picture	.768	.002
	Concrete_no_picture	.573	.119
Abstract_no_picture	Abstract_picture	.470	.480
	Concrete_picture	.747	.000
	Concrete_no_picture	.551	.001
Concrete_picture	Abstract_picture	.768	.002
	Abstract_no_picture	.747	.000
	Concrete_no_picture	.681	.128
Concrete_no_picture	Abstract_picture	.573	.119
	Abstract_no_picture	.551	.001
	Concrete_picture	.681	.128

© Language Learning & Technology

Results: Paired Samples T-test

- Are high-imagery words ‘easier’ to acquire than low-imagery words?

		Mean	N	Std. Deviation	Std. Error Mean	t	Sig,
Post-test	Total_abstract	17.92	50	9.044	1.279	4.731	.000
	Total_concrete	22.98	50	5.549	.785		
Delayed post-test	Total_abstract	10.36	50	9.264	1.310	4.987	.000
	Total_concrete	15.58	50	8.104	1.146		

© Language Learning & Technology

Interpreting results

- To what extent do the data answer the research question(s)?
  - What reliable conclusions can be drawn?
    - scope, applicability, etc.
- To what extent do the results support previous findings?
- To what extent does the study contribute to existing knowledge in the field?

© Language Learning & Technology

Limitations

- Study design
  - Methodology: variables
    - choice of test items
    - duration of study
    - type of data collection
    - …
  - Study participants
    - variables, e.g., L1,…

© Language Learning & Technology

Descriptive Statistics

- **Distribution** lists every value of a variable and the number of subjects who had each value.

- **Central tendency** is an estimate of the "center" of a distribution.
  - **Mean** is the most commonly used method . It assumes a normal distribution of values.
  - **Median** is the score at the exact middle of the set of values. It makes no assumption of normalcy.
  - **Mode** is the most frequent value in a set of scores.

- **Dispersion** is the spread of the values around the central tendency.
  - **Range** is the highest minus the lowest score. It is easily influenced by outliers.
  - **Standard deviation** is the relationship of a set of scores to the sample mean. It is a more accurate estimate of dispersion because it is less influenced by outliers. It assumes a normal distribution of values.

Research Methods Knowledge Base

© Language Learning & Technology

Inferential statistics: T-test

The t-test is appropriate whenever you want to compare the means of two groups, and especially appropriate for the posttest-only two-group designs

- **Paired or unpaired**

One-to-one correspondence between the values in the two samples, i.e., X_1 corresponds to Y_1, etc. The formulas for paired data are simpler than for unpaired

- **Equal or unequal variances**

Equal variances yield somewhat simpler formulas, although with computers this is no longer a significant issue

Quick T-test calculator

© Language Learning & Technology

Inferential statistics: ANOVA

- One-way ANOVA is typically used to test for differences among **at least** three groups, since the two-group case can be covered by a T-test. Multiple t-tests are inappropriate because they lead to inflation of Type I error (false positive).
- One-way ANOVA for repeated measures is used when repeated treatments are used with the same subjects.
- Two-way ANOVA is used when there are two independent variables.

© Language Learning & Technology

Inferential statistics: MANOVA & ANCOVA

- **Multivariate Analysis of Variance (MANOVA)** is a type of ANOVA with several dependent variables. Commercial package SPSS MANOVA

- **Analysis of Covariance (ANCOVA)** is used for designs with both pre- and posttests. ANCOVA removes the effect of the pretest so that you can just look at the difference in the posttest measurements between the treatment and comparison groups. Some versatile point & click commercial packages that do ANCOVA are MINITAB and SPSS. There are also easy-to-use online aids, such as http://faculty.vassar.edu/lowry/vsancova.html.

© Language Learning & Technology

Statistics on youtube.com

- Basics of quantitative and qualitative research methods
- Mixed research methods
- Introduction to statistics (SPSS)
  - These are several video clips, each covering different aspects of statistical analyses performed in SPSS.
- Changing SPSS statistics to APA style
- Exporting SPSS output to Word

© Language Learning & Technology

Part 5

Preparing to publish

© Language Learning & Technology

Preparing to publish

- Follow carefully the requirements of the journal to which you submit your manuscript.
  - For example, when submitting to *LLT*, consult the Information for Contributors.

- From the beginning, pay attention to style, word limit, inclusion of hyperlinks, multimedia.
  - "Style" refers to headings, sub-headings, citations within the body of the text, format for references.

© Language Learning & Technology

Writing the article

- Writing the **abstract**:
  - May be the most important paragraph in your article (though the last item you write).
  - It is a brief, comprehensive summary.
  - Make each sentence maximally informative.
  - Begin with the most important information, usually the purpose or thesis.
  - Then briefly describe the key elements of the study.
  - State the results and conclusions.

© Language Learning & Technology

Editing the article

- Follow the journal's required style sheet.
- Seek assistance with editing as needed:
  - ask a colleague to read for content
  - ask for help with English academic prose if you are not a native English speaker/writer.
  - proofread carefully! Use a spell-checker!
  - double-check that the References list is complete.

© Language Learning & Technology

Submitting your article

- Submit to only one journal (at a time).
- Be prepared to use an online submission system, such as ScholarOne (formerly Manuscript Central).
- You will also need to get used to receiving “automated” e-mail responses (do not ignore them!).
- Inform yourself about the usual review procedure for the journal.

© Language Learning & Technology

Receiving a decision

- If you have not received a decision on your submission, e-mail the editor(s) only after a suitable period of time has elapsed (check the journal's average time from submission to decision).
- If your article is not "accepted as is" or "accepted with minor revisions" (both are rare!), read the reviews, then set them aside for a few days.

© Language Learning & Technology

Revising your article

- If the decision is “Revise and Resubmit,” after waiting a few days, carefully re-read the reviewers’ comments; they probably won’t look so bad ☺.
- If the editors invite you to revise and resubmit your article, carefully follow the reviewers’ suggestions and note points that you do not agree with or decide not to follow.
- Your revised paper will inevitably be improved and stronger.

© Language Learning & Technology

Resubmitting your article

- Write a cover letter to the editors.
- List all of the changes and additions you made.
- Discuss any issues that were suggested by the reviewers or the editors that you did not agree with.
- Do not be discouraged if you are asked to revise your manuscript again. Chances are that the editors think your manuscript is publishable and that further revisions will improve your article. They are usually right about the latter ☺.

© Language Learning & Technology

Questions?

Thank you!

© Language Learning & Technology