Chapter 8

How to Conduct Experimental Research Illuminating Gesture's Role in L2 Acquisition

Laura M. Morett, University of Alabama

Abstract

Experimentation provides a powerful methodological tool for investigating gesture's role in second language (L2) acquisition because it allows the effect of gesture on various aspects of L2 acquisition to be isolated. Previous experimental research has demonstrated that the information that gesture conveys in relation to speech determines gesture's impact on L2 acquisition (Brown & Gullberg, 2008; Morett, 2014; Morett & Chang, 2015; Xi et al., 2020; Zhang et al., 2020). Using previous experimental research as a guide, this chapter discusses how experimentation has been fruitfully employed to illuminate gesture's role in L2 acquisition. It discusses how to develop and conduct well-designed experiments capable of yielding insight into gesture's role in L2 acquisition and how to analyze data resulting from them, providing readers with the conceptual tools necessary to conduct experimental research illuminating gesture's role in L2 acquisition.

Keywords: Experimental research, second language acquisition (SLA), gesture

Introduction

Experimentation is unique methodologically because it permits causal conclusions to be reached and alternative explanations to be ruled out. It involves testing a hypothesis via systematic manipulation of one or more independent variables, such as presence or type of gesture, to determine their effect on one or more dependent variables, such as L2 vocabulary recall or speech sound discrimination. Therefore, it allows researchers to isolate the effect of gesture on various aspects of L2 acquisition. Experimentation has provided insight into gesture's role in many aspects of L2 acquisition, including vocabulary learning (Allen, 1995; Andrä et al., 2020; Huang et al., 2019; Kelly et al., 2009; Morett, 2014, 2018, 2019; Porter, 2016; Tellier, 2008), L2 speech sound acquisition (Amand & Touhami, 2016; Baills et al., 2019; Gluhareva & Prieto, 2017; Hoetjes & Van Maastricht, 2020; Morett & Chang, 2015; Xi et al., 2020; Zhang et al., 2020; Zhen et al., 2019; Zheng et al., 2018), and speech and listening comprehension (Dahl & Ludvigsen, 2014; Drijvers & Özyürek, 2020; Sueyoshi & Hardison, 2005; Wagner, 2010). Thus, when used appropriately, experimentation is a powerful methodological tool for investigating gesture's role in L2 acquisition.

In the field of gesture studies, experimentation is a relatively recent methodological approach. Indeed, seminal gesture research, much of which adopted a cross-linguistic and cross-cultural perspective, exclusively employed systematic observational and descriptive approaches (Efron, 1941; Kendon, 2004). Such approaches, which are typically applied to corpora (databases of transcribed language), continue to inform the understanding of gesture's role in L2 acquisition (Kimura & Kazik, 2017; Lazaraton, 2004; Smotrova & Lantolf, 2013; van Compernolle & Smotrova, 2017; Wang & Loewen, 2016). Since David McNeill's seminal theoretical and taxonomical work relating gesture to speech (McNeill, 1985; McNeill & Levy, 1982), however,

there has been an increase in experimentation as a methodological approach in gesture studies, including in the content area of L2 acquisition.

In the next section of this chapter, I provide a theoretical background concerning how experimentation has been fruitfully employed to illuminate gesture's role in L2 acquisition. Subsequently, I will discuss how to develop and conduct well-designed experiments capable of yielding insight into gesture's role in L2 acquisition and how to analyze data resulting from them.

Theoretical Background

This section of the chapter discusses how experimental research has illuminated gesture's role in several key aspects of L2 acquisition: vocabulary learning, L2 speech sound acquisition, speech and listening comprehension, corrective feedback, and speech production.

Vocabulary Learning

To date, the aspect of L2 acquisition into which experimentation incorporating gesture has provided the greatest insight is vocabulary learning. In a seminal experiment examining gesture's effect on L2 vocabulary acquisition (Allen, 1995), English-speaking postsecondary students in a first-semester French course learned novel French spoken expressions in two conditions: one viewing and producing emblematic gestures related to their meaning or not doing so. Immediately after learning and at several subsequent intervals, participants heard these expressions and wrote down their English meanings. Learning French expressions with gestures made students more likely to recall their English meanings on posttests than learning French expressions without gestures, demonstrating that learning L2 vocabulary with semantically related gestures enhances memory for its L1 meanings.

More recently, there has been an increase in experimental research elucidating gesture's impact on L2 vocabulary learning (see Macedonia, 2014, for a review). This research has revealed that gestures are at least as effective as pictures in promoting learning when they convey the meanings of L2 vocabulary (Andrä et al., 2020; Morett, 2019; Porter, 2016; Repetto et al., 2017; Tellier, 2008), that gestures promote learning when they are not incongruent with the meanings of L2 vocabulary (Garcia-Gamez & Macizo, 2019; Gullberg et al., 2012; Huang et al., 2019; Kelly et al., 2009; Kushch et al., 2018; Macedonia et al., 2011; Macedonia & Knösche, 2011; Mavilidi et al., 2015; Mayer et al., 2015; Morett, 2014), and that gesture production is at least as effective as gesture observation in promoting L2 vocabulary learning (Macedonia et al., 2014; Morett, 2018; Sweller et al., 2020). However, this research has also revealed that gestures conveying the meanings of L2 vocabulary may be ineffective learning aids for phonologically similar words (Kelly & Lee, 2012; Morett & Chang, 2015), and such gestures may not always promote L2 vocabulary learning even for phonologically distinct words (Krönke et al., 2013; Rowe et al., 2013). Taken together, this experimental work indicates that gestures conveying the meanings of L2 words enhance their acquisition, especially when those words are phonologically distinct and when gestures are produced (see Study Box 1).

[INSERT STUDY BOX 1 ABOUT HERE]

These experimental findings complement findings concerning gesture's role in L2 vocabulary learning derived from non-experimental research, including case studies, which entail detailed examination of the behavior of a single individual or dyad within a real-world context (Lazaraton, 2004; McCafferty, 2002; van Compernolle & Smotrova, 2017) and qualitative research, which entails first-hand observation of participants in classroom or real-world settings (McCafferty, 1998; Smotrova & Lantolf, 2013). For example, experimental findings

demonstrating that gestures conveying the meanings of novel L2 words promote their acquisition relative to the absence of such gestures (Allen, 1995; Kelly et al., 2009; Macedonia & Knösche, 2011) complement non-experimental findings showing that L2 instructors and learners often produce such gestures in conjunction with novel L2 words (Lazaraton, 2004; McCafferty, 2002; Smotrova & Lantolf, 2013). Thus, experimentation is informative both alone and in combination with non-experimental research in revealing gesture's role in L2 vocabulary learning.

L2 Speech Sound Acquisition

Another aspect of L2 acquisition into which experimentation incorporating gesture has provided substantial insight is L2 speech sound acquisition. Some of the first experiments examining gesture's effect on the acquisition of L2 segmental phonology suggest that gestures conveying Japanese vowel length contrasts have little or no effect on their perception by L1 English speakers (Hirata et al., 2014; Hirata & Kelly, 2010; Kelly et al., 2014, 2017) although other experiments in this vein suggest that such gestures enhance production of Japanese vowel length contrasts by L1 English and Catalan speakers (Iizuka et al., 2020; Li et al., 2020; Roberge et al., 1996). Moreover, subsequent experiments examining gesture's effect on the acquisition of other L2 phonemes demonstrate that gestures conveying pitch contours of lexical tones facilitate their perception and production by L1 English-speakers in L2 Mandarin (Baills et al., 2019; Hannah et al., 2017; Morett & Chang, 2015; Zhen et al., 2019; Zheng et al., 2018), gestures mimicking articulation of L2 phonemes facilitate their perception and production (Amand & Touhami, 2016; Hoetjes & Van Maastricht, 2020; Xi et al., 2020), and hand clapping facilitates perception and production of L2 Japanese segmentals (Iizuka et al., 2020). These findings provide evidence that, when encountered in conjunction with L2 phonemes, many gestures conveying relevant acoustic characteristics visually facilitate their acquisition (see Study Box 2).

[INSERT STUDY BOX 2 ABOUT HERE]

In addition to L2 phonemes, experimentation also provides evidence that gestures facilitate acquisition of L2 suprasegmentals. Indeed, recent experimental research indicates that production of hand clapping as well as observation and production of beat gestures—simple rhythmic gestures conveying prominence—in conjunction with L2 speech improve production of isochrony (Gluhareva & Prieto, 2017; Llanes-Coromina et al., 2018; Zhang et al., 2020). In addition, it shows that observing gestures conveying L2 intonational pitch contours improves L2 intonational production (Yuan et al., 2019), paralleling similar findings for L2 lexical tone acquisition (Baills et al., 2019; Zheng et al., 2018). Together, these findings demonstrate that when encountered in conjunction with L2 prosody, gestures conveying relevant acoustic characteristics visually enhances its acquisition, similar to their effect on L2 phonemes.

As is the case for gesture's role in L2 vocabulary acquisition, these experimental findings complement non-experimental findings derived from case studies and qualitative research concerning gesture's impact on L2 speech sound acquisition. For example, experimental findings demonstrating that gestures conveying L2 isochrony via temporal synchronization with speech improves L2 pronunciation (Gluhareva & Prieto, 2017; Llanes-Coromina et al., 2018; Zhang et al., 2020) complement non-experimental findings showing that L2 instructors and learners often produce rhythmic body movements in conjunction with L2 speech to convey L2 isochrony (Hudson, 2011; McCafferty, 2006; Smotrova, 2017). Thus, as is the case for L2 vocabulary learning, experimentation is informative both alone and in combination with non-experimental research in revealing gesture's role in L2 speech sound acquisition.

Speech and Listening Comprehension

A third aspect of L2 acquisition into which experimentation incorporating gesture has provided insight is speech and listening comprehension. In a seminal experiment examining gesture's impact on comprehension of L2 spoken discourse (Sueyoshi & Hardison, 2005), low-intermediate and advanced postsecondary L2 learners of English viewed a recorded lecture on an unfamiliar topic in English in which the speaker was either visible and produced gestures, visible and did not produce gestures, or not visible. Subsequently, participants answered questions concerning their comprehension of the content of the lecture. Comprehension of participants with low-intermediate L2 proficiency was highest when the speaker was visible and produced gestures. By contrast, comprehension of participants with advanced L2 proficiency was highest when the speaker was visible and did not produce gestures (see Study Box 3). The finding that gesture enhanced L2 listening comprehension in lower proficiency listeners is consistent with the results of a previous study demonstrating that elementary school children learning English as an L2 comprehended English stories better when they were told by a storyteller who gestured and produced repetitions and comprehension checks than when they were told by a storyteller who did not make these interactional adjustments (Cabrera & Martínez, 2001). Thus, these findings suggest that comprehension of L2 spoken discourse may be enhanced by observation of accompanying gestures in listeners with low but not high L2 proficiency.

INSERT STUDY BOX 3 HERE

In a subsequent study in a similar vein comparing gesture's impact on spoken discourse comprehension by L1 and L2 English speakers (Dahl & Ludvigsen, 2014), middle school students watched videos of a speaker describing non-visible cartoon images while either gesturing or not gesturing and then drew pictures of each image described. In contrast with L1

English speakers, who drew equally accurate depictions regardless of whether they had viewed videos with gesture, L2 English speakers who had viewed videos with gesture drew more accurate depictions than those who had viewed videos without gesture. In another study on this topic (Wagner, 2010), L2 learners of English varying in proficiency took a test of L2 listening comprehension in which English discourse was presented either via audio only or audio accompanied by video of a speaker producing it in a semi-scripted manner, gesturing in some cases. Participants who took the audiovisual test demonstrated performance on it superior to that of participants who took the audio only test. Item analysis revealed that compared to participants who took the audio only test, participants who took the audiovisual test were more likely to provide correct responses to items for which gesture provided pertinent information. In contrast, in a study examining gesture's impact on L2 speech perception (Drijvers & Özyürek, 2020), L1 Dutch listeners and German listeners highly proficient in Dutch viewed videos of a speaker producing Dutch verbs in clear, moderately degraded, or severely degraded speech with lips blurred or visible. When the lips were visible, verbs were either accompanied by gestures conveying their meanings or unaccompanied by gestures. Compared to L1 listeners, L2 listeners were less likely to benefit from combined enhancement of visible speech and gestures, particularly because the benefit from visible cues was minimal when speech quality was degraded. Taken together, the findings of the first two experiments suggest that gesture enhances discourse comprehension in L2 learners to a greater degree than it does in L1 speakers, whereas the findings of the third suggest that this may not be the case for speech perception.

Compared to L2 vocabulary learning and speech sound acquisition, only one non-experimental study has examined gesture's impact on L2 speech and listening comprehension (Kida, 2008). Nevertheless, its findings complement those of the experimental work discussed

above by demonstrating that observing gestures in conjunction with spoken L2 discourse enhances comprehension of it. Together, this finding and the paucity of non-experimental research on the topic demonstrate that experimentation is a crucial methodological tool for illuminating gesture's role in L2 speech and listening comprehension.

Corrective Feedback and Speech Production

Aside from vocabulary learning, phonological acquisition, and speech and listening comprehension, experimental research has examined gesture's role in L2 corrective feedback and speech production. With respect to corrective feedback, experimentation has revealed that, relative to no feedback and recasts only, gestures that convey locative prepositions during recasts of them enhance low-intermediate proficiency L2 English learners' correct production of them following a week-long delay (Nakatsukasa, 2016). In contrast, point-back gestures conveying the past accompanying recasts of regular past tense verbs fail to enhance the same population's correct production of them (Nakatsukasa, 2019). Together, these findings suggest that gestures accompanying corrective feedback may be more effective for item-based rather than rule-based learning.

With respect to gesture and speech production, experimentation has revealed that L2 speakers produce overexplicit gestures regardless of whether they are speaking to a visible or non-visible interlocutor (Gullberg, 2006) and that deictic gesture production when speaking to visible vs. non-visible interlocutors positively predicts their recall of L2 vocabulary (Morett, 2018). Moreover, research has revealed that gesture restriction hinders bilinguals' storytelling in their L2 but not their L1 (Laurent et al., 2015). These findings suggest that gesture production facilitates L2 speech production, particularly when interlocutors are visible.

Important Considerations

There are a number of considerations that you should take into account to successfully design and implement an L2 gesture experiment permitting justifiable causal conclusions.

First, you will need to determine the purpose of your experiment in relation to previous research. What theory or research question will your experiment address? Will your experiment replicate and seek to reproduce the findings of a previous experiment? What gap in the literature will your experiment fill? How will your experiment build upon previous research? What are your hypotheses, or predictions, for your experiment? These are all important questions that will influence the specifics of your experiment, so you should be able to answer all of them prior to proceeding.

Second, you should define the setting and participants of your experiment. Typical settings for such experiments include the laboratory, the internet, and the classroom. These settings may constrain the type and number of participants you can recruit. For example, a classroom experiment will necessarily be limited to students enrolled in targeted classes, whereas internet-based experiments can be completed by participants anywhere in the world. Furthermore, classroom experiments are typically conducted on matched groups of participants, making them quasi-experiments, whereas lab- and internet-based experiments are typically conducted on individuals. Also, the data you can collect will be affected. For example, response time data can be collected in the lab and on the internet but not in the classroom.

Third, in designing your experiment, you should think about what your experimental and control (baseline) conditions are and what their duration is. In experimental research examining gesture's impact on L2 acquisition, the experimental conditions often include L2 speech with gesture, whereas the control condition often includes L2 speech without gesture, and the duration

should be long enough to affect performance. You should also think about what data (dependent variables) will be collected and when and how they will be collected. The answer to this question depends upon the objective of the research as well as its setting. For example, if your experiment examines gesture's impact on L2 word learning in the lab, you could use specialized software or computer code to measure participants' recall and/or recognition of the meanings of L2 words as well as the amount of time taken to recall and/or recognize each L2 word via button presses both immediately after learning and after a delay (e.g., of one week) has elapsed. In contrast, if your experiment examines gesture's impact on L2 word learning in the classroom, you could measure recall and/or recognition of the meanings of L2 words by having participants write down or mark their L1 meanings using a sheet of paper both immediately after learning and after a delay (e.g., of one week) has elapsed, but it is not possible to measure the amount of time required by each participant to recall and/or recognize each L2 word.

Fourth, you should decide upon the stimuli that you will use in your experiment. What aspect(s) of the L2 will you examine gesture in relation to? Will your stimuli contain or elicit gesture, and if so, what type(s)? Can you obtain your stimuli from a public database, or will you need to construct your own? If you will need to construct your own stimuli, you may need to conduct a norming study prior to your experiment to ensure that participants interpret them in accordance with your expectations. These norming results are often helpful in interpreting experimental results, particularly when they are inconsistent with the hypotheses. Additionally, they are sometimes requested by reviewers as part of grant applications and manuscripts.

Fifth, you should develop and pilot the procedure of your experiment. To do so, you should construct a list of actions to be executed by the experimenter and script any verbal instructions to standardize them for all participants. You should then test this procedure with a

small number of participants to determine if any changes need to be made. If changes are necessary, the revised procedure should be piloted again with new participants. If no changes are necessary, data collected from these participants can be combined with data collected from participants subsequently using the same procedure.

Finally, you should consider how you will analyze the data from your experiment. If your experiment involves analyzing gestures produced by participants, you should determine what aspects of these gestures are relevant (frequency, type, size, etc.), obtain training in gesture coding by attending a workshop or consulting with a researcher with expertise in gesture analysis, and have another gesture researcher independently code a portion of your data to ensure inter-rater reliability. You should consider what statistical tests you will apply to your data and familiarize yourself with their assumptions and how to execute them.

Like other methods used to study gesture's impact on L2 acquisition, experimentation has advantages and disadvantages that should be carefully considered before deciding whether it is the most appropriate method to address a research question. The greatest advantage of experimentation is that it provides high internal validity by permitting control over all (or most) aspects of a study, allowing conclusions about causality to be reached and alternative explanations to be ruled out. In experimental research, the experimenter determines the experimental design, the independent and dependent variables, the experimental task, the stimuli, the data, the participant sample, and how the experiment is administered. By making informed decisions about these aspects of an experiment, it is possible to minimize confounds (variables influencing both the independent and dependent variables) that may influence the results. By the same token, the greatest disadvantage of experimentation is that it provides low external validity, limiting generalization of the results to contexts differing from those of the experiment. Thus,

caution should be exercised when considering the ramifications of laboratory and internet experiments examining gesture's role in L2 acquisition for pedagogical practice.

Purpose of Study

Experimentation entails testing a hypothesis via systematic manipulation of one or more independent variables, such as presence or type of gesture, to determine their effect on one or more dependent variables, such as L2 vocabulary recall or speech sound discrimination. For example, as discussed above, experimentation is used to examine how gestures conveying different information affect various aspects of L2 acquisition, such as vocabulary learning, speech sound acquisition, and speech and listening comprehension (e.g., Morett & Chang, 2015; Sueyoshi & Hardison, 2005; Xi et al., 2020). It is also used to examine how manipulating relevant aspects of the communicative environment (e.g., interlocutor variability, hand movement restriction) affects gesture and L2 speech production in L2 learners (Gullberg, 2006; Laurent et al., 2015; Morett, 2014). Ideally, experimentation should be leveraged to test the predictions of theories concerning gesture's role in L2 acquisition (see Gullberg, 2006, for examples). Experimentation would be inappropriate and is therefore not used to investigate descriptive research questions such as how L2 instructors gesture in the classroom (Smotrova & Lantolf, 2013) and how L2 learners gesture when producing the target language (Eskildsen & Wagner, 2013) as well as associational research questions such as how gesture production is associated with L2 fluency (Graziano & Gullberg, 2013).

Research Questions

Typical research questions for experiments examining gesture's role in L2 acquisition include the following:

- Does observing gestures conveying referents facilitate L2 word learning to a greater degree than observing images conveying referents (Andrä et al., 2020; Morett, 2019; Tellier, 2008)?
- How does the information conveyed via gestures affect L2 acquisition?
    - How does it affect perception (Hirata et al., 2014; Hirata & Kelly, 2010; Kelly et al., 2017; Morett & Chang, 2015; Zhen et al., 2019) and production (Xi et al., 2020; Zheng et al., 2018) of L2 phonemes?
    - How does it affect production of L2 suprasegmentals (e.g., isochrony, intonation; Baills et al., 2019; Gluhareva & Prieto, 2017; Llanes-Coromina et al., 2018; Yuan et al., 2019; Zhang et al., 2020; Zheng et al., 2018)?
    - How does it affect L2 vocabulary learning (Kelly et al., 2009; Morett, 2014, 2018; Morett & Chang, 2015), particularly with respect to characteristics of words such as phonological similarity (Kelly & Lee, 2012) or word class (Garcia-Gamez & Macizo, 2019)?
    - How does it affect L2 (morpho)syntax acquisition via recasts (Nakatsukasa, 2016, 2019)?
- Does observing gestures conveying referents enhance L2 speech and listening comprehension (Dahl & Ludvigsen, 2014; Drijvers & Özyürek, 2020; Sueyoshi & Hardison, 2005; Wagner, 2010)?
- How does manipulation of relevant aspects of the communicative environment (e.g., interlocutor variability, hand movement restriction) affect gesture and L2 speech production (Gullberg, 2006; Laurent et al., 2015; Morett, 2014)?

Methods

This section provides you with some guidelines for designing an experiment examining gesture's role in L2 acquisition.

Participants

Generally, the number of participants to be recruited should be determined prior to data collection to avoid under- or over-sampling, which may result in statistically unsubstantiated conclusions or unnecessary time and effort expenditure. Ideally, a statistical technique known as power analysis should be used to determine the appropriate sample size based on several of the experimental design characteristics discussed subsequently in this section as well as the anticipated strength of the results (i.e., the difference between conditions, such as gesture vs. none). Alternatively, sample sizes may be determined based on those of previous experiments with similar methods and designs. Determination of sample sizes may take into account practical considerations such as the project timeline and participant availability, as well. As a rule of thumb, you should aim to recruit at least 15 participants per condition.

Demographic Data

It is good practice to collect data concerning participants' demographic characteristics. At minimum, these characteristics include age and gender, and they may also include socioeconomic status and significant locations (birthplace, hometown, place(s) of residence), given that all of these characteristics may influence L2 acquisition. Typically, demographic data is collected via self-report measures such as questionnaires, which are often administered at the beginning or end of experiments. These data are generally either descriptive or are accounted for in analysis as covariates (see Data Analysis section) but may also be used as a screening variable to ensure that all participants are of a similar age, for example.

Language Exposure and Proficiency

In L2 acquisition research, it is important to consider participants' exposure to and proficiency in the L2 as well as their L1 and any additional languages to understand how it may affect their acquisition of the target language. Considerations with respect to these languages may include, but are not limited to, age of acquisition; duration, frequency, nature, and modality of exposure; bilingualism; formal language education; and mastery of specific aspects of them (e.g., L2 phonemes, vocabulary, specific syntactic structures). Language exposure is typically ascertained via standardized or experimenter-created self-report inventories such as the Language Exposure and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). In contrast, language proficiency, including mastery of specific aspects of languages, is often measured directly via standardized assessments (e.g., DLI, TOEFL, PPVT; Dunn & Dunn, 2007).

Materials

Most experiments are conducted using computers, which are used to control the timing of presentation of materials or stimuli and to record response data from participants. When considering stimuli to include in an experiment, it is crucial to have determined the independent and dependent variables. As the names suggest, manipulation of independent variables isn't affected by dependent variables, whereas dependent variables may vary based on manipulation of independent variables. The independent variables comprise at least two conditions, or levels, based on their presence and characteristics: experimental and control. The dependent variables are quantitative and may be either categorical, such as whether a word was recalled correctly or not, or continuous, such as the number of words recalled correctly over the course of an experiment. Both variables represent operationalizations of the constructs that the experiment is designed to probe, whether they are an aspect of L2 acquisition (e.g., lexical tone, grammatical

aspect, oral fluency) or gesture (e.g., frequency, type, form). When multiple independent or dependent variables are present in an experiment, any overlap between variables of each type should be accounted for in data analysis (see Data Analysis section).

Pre- and Post-Tests

Pre-tests are often conducted to determine whether participants have pre-existing knowledge of the constructs that are the focus of the experiment. Pre-tests precede the experimental task and probe knowledge of these constructs directly in terms of the dependent variables as opposed to relying on self-report measures. Pre-tests should ideally be as similar to post-tests as possible but not exactly the same as the post-test. For example, a pre-test for an experiment concerning gesture's impact on acquisition of L2 phonemes might consist of perception and production of those phonemes in L2 words. If pre-tests are expected to affect post-test performance or participants are unlikely to have any knowledge of factors that are the focus of the experiment, pre-tests should not be used.

Post-tests probe participants' knowledge of constructs that are the focus of the experiment following the experimental task. As such, they measure the extent to which participants' knowledge of these constructs reflects mastery and are therefore an essential component of experiments probing L2 acquisition. For example, a post-test for an experiment concerning gesture's impact on acquisition of L2 phonemes might consist of perception and production of those phonemes in a different set of L2 words than the pre-test. Comparison of the results of post-tests and similar pre-tests reflects how well these factors have been learned in the experiment.

Procedures

Experimental procedures should facilitate experiment completion. They should be standardized to minimize variation between participants; thus, it is generally advisable to script them prior to implementing an experiment. For example, all participants need to sign an informed consent form approved by the Institutional Review Board of the host institution. Then, if a pre-test, treatment, post-test design is used, they need to complete the pre-test measurement, participate in the training, and complete the post-test measurement. Multiple versions of paradigms are typically necessary to implement systematic manipulation of independent variables and counterbalancing. These versions should be evenly distributed among participants via random assignment to the extent possible. For example, an experiment with a within-participants design in which some L2 words are paired with gestures and others are not should have two versions to ensure that each L2 word is presented to half of the participants with gesture and the other half of the participants without gesture. Likewise, in a version of this experiment with a between-participants design, all L2 words should be paired with gesture for half of the participants and should be unaccompanied by gesture for the other half of the participants.

Piloting

The goal of piloting is to ensure that the experiment runs smoothly with participants representative of the target population. Piloting typically involves testing experimental paradigms and procedures on a small number of eligible participants while attending closely to their ability to complete them successfully. If any problems such as experimental design attributes, participant confusion, or experiment malfunctions are identified in piloting, an attempt to rectify them should be made, and rectification should be verified with additional pilot

participants. This process is repeated iteratively until no further problems are identified. Pilot data collected when any problems are identified should be discarded, but pilot data collected when no problems are identified can be included in analyses.

Data Collection

Once experimental procedures and paradigms are developed and piloted, full-scale data collection can commence. Data can be collected in the laboratory, classroom, or via internet-based experimentation platforms. Laboratory and classroom data collection allows the experimenter to monitor participants during the experiment and provides greater control over the ambient environment (e.g., lighting, noise), whereas internet-based experimentation permits quicker and more convenient data collection, including greater access to participants with diverse characteristics of interest. Data should be collected from a pre-determined number of participants (see Participants section) using standardized methods. At this point, changes to experimental procedures and paradigms should only be made if a critical error is discovered, and data collected prior to any changes should be discarded and re-collected.

Transcription, Coding, and Identifying of Relevant Behavior

If your experiment examines speech production, all speech produced by participants within it will need to be transcribed. You should determine the level of detail to include in the transcription by considering your research question and dependent variables of interest. For example, if your experiment investigates gesture's impact on production of L2 phonemes within words, you will need to transcribe L2 words phonetically using a convention such as the International Phonetic Alphabet. If your experiment investigates gesture's impact on L2 word-meaning association via backwards translation from L2 to L1, however, use of such conventions are unnecessary.

If your experiment examines gesture or speech production, you will need to develop a coding scheme capturing the dependent variables of interest. If you are interested only in the number of gestures produced by participants without regard to their timing or the information that they convey, your coding scheme can consist simply of coding all meaningful hand (or body) movements. If you are interested in the timing of gestures and/or the information that they convey, however, your coding scheme should reflect this by locating gestures relative to each other and speech temporally and by indicating relevant information such as type (e.g., iconic, beat, deictic) or redundancy with co-occurring speech (i.e., same vs. different information). Similar guidelines apply to coding speech. Thus, if you are interested only in whether L2 words are pronounced or associated with L1 glosses correctly, your coding scheme can be categorical (e.g., 0 = incorrect; 1 = correct). If you are interested in how native-like pronunciation of L2 phonemes is, however, your coding scheme should be continuous (e.g., using 1-7 scale).

Data Analysis

Once the data have been collected, they can be subjected to statistical analyses. In the remainder of this section, I provide a brief, non-technical overview of some of the most common statistical analyses used to analyze experimental data concerning gesture's role in L2 acquisition.

T-Test

T-tests are used to determine whether a continuous dependent variable differs significantly between two levels of a categorical independent variable. Thus, they can be used to analyze data from experiments in which one or more independent variables with two levels are manipulated that produce continuous dependent variables. For example, they could be used to examine recall of a set of L2 vocabulary words learned with vs. without gesture. T-tests come in two varieties: *independent samples*, in which the independent variable varies between

participants, and *paired samples*, in which the independent variable varies within participants. Both varieties of t-tests assume that the data are distributed in a bell-shaped curve (normality) and that the dependent variables vary to a similar degree for both levels of the independent variable (homoscedasticity). Because t-tests are limited to independent variables with two levels, other analyses must be used to analyze similar data from experiments with multiple independent variables and independent variables with more than two levels.

ANOVA, MANOVA, and ANCOVA

Analysis of variance (ANOVA) is used to determine whether a continuous dependent variable differs significantly between categorical independent variables, at least one of which has more than two levels. Similar to t-tests, independent variables in ANOVA can vary either between participants or within participants, the latter of which is referred to as *repeated measures*. ANOVA has several advantages over t-tests, including flexibility concerning the number of levels in independent variables; the ability to examine the effects of multiple independent variables, referred to as multivariate analysis of variance, or MANOVA, both separately (main effects) and conjointly (interactions); and the ability to account for differences in continuous variables that are not the focus of the analysis but may affect the dependent variable (covariates), referred to as analysis of covariance, or ANCOVA. If only one independent variable with two levels is present, however, t-tests should be used to analyze the data as they are more parsimonious than ANOVA. Within the context of ANOVA used to analyze data from psycholinguistic experiments, independent variables are referred to as *fixed factors* because they are systematically manipulated, whereas participants and trials are referred to as *random factors* because they represent a representative sample reflecting random variation in irrelevant characteristics taken from a larger population. ANOVA assumes that the data are distributed in a

bell-shaped curve (normality) and that the dependent variables vary to a similar degree for both levels of the independent variable (homoscedasticity), and they assume that observations are independent of one another (i.e., that they aren't highly correlated; independence of observations). Disadvantages of ANOVA include more assumptions than some other types of analyses; restriction to data with categorical independent variables and continuous dependent variables; and restriction to a single random factor per analysis, necessitating separate analyses by participants and trials.

Hierarchical Linear Models

Hierarchical linear models (HLMs) are used to examine how dependent variables differ by independent variables that vary at more than one level. Within the context of psycholinguistic experiments, this nested variation is typically present in random factors (i.e., trials, participants), allowing a single HLM to take variation in both of these factors into account. HLM has the same assumptions as linear regression and ANOVA—linearity, normality, homoscedasticity, and independence of observations—some of which are modified for hierarchical data. Similar to ANOVA, independent variables in HLMs may vary either within or between participants, there are no restrictions on the number or levels of independent variables, and covariates can be included in analyses. Unlike all of the other analyses discussed in this section, however, all variables in HLMs may be either categorical or continuous; thus, HLMs are commonly referred to as *linear mixed-effect models* in psycholinguistic research. Application of HLMs to analysis of experimental psycholinguistic data, including data concerning gesture's impact on L2 acquisition, is relatively recent and growing due to their flexibility.

Future Research Ideas

Although experimentation has revealed many important aspects of gesture's role in L2 acquisition, it has the potential to reveal even more. At present, there is a pressing need to examine how gesture observation affects aspects of L2 acquisition other than vocabulary learning, L2 speech sound acquisition, and speech and listening comprehension. For example, the only experimental research to date that has examined gesture's impact on L2 (morpho)syntax acquisition has done so via corrective feedback (Nakatsukasa, 2016, 2019), so experimentation examining it via initial instruction of L2 (morpho)syntactic forms is needed. Moreover, few experiments to date have examined gesture production in relation to L2 acquisition (Gullberg, 2006; Laurent et al., 2015; Morett, 2014), so more experimental research on this topic is needed, particularly with respect to whether instructing L2 speakers to produce gesture while speaking increases the quantity and quality of their L2 speech production.

In addition, there is a need for experimental research comparing the effects of different types of gestures on various aspects of L2 acquisition. To date, no experimental research has examined how different types of gestures affect L2 listening comprehension, so it is currently unclear whether—and how—their impact on it differs. Relatedly, although experimental research has examined how gestures conveying referents vs. emphasis affects memory for key words and phrases in L1 (So et al., 2012) and the effects of each of these types of gestures vs. no gesture on it independently in L2 (Garcia-Gamez & Macizo, 2019; Kelly et al., 2009; Kushch et al., 2018), no experimental research to date has compared the effects of these types of gestures on word memory in L2.

Finally, there is a need to examine whether gesture's impact on various aspects of L2 acquisition differs based on L2 proficiency. To date, most experimental research has examined

gesture's role in L2 acquisition in learners with little or no knowledge of the target language. Thus, these experiments should be replicated with more proficient L2 learners to determine whether the results differ in this population, and if so, how. Likewise, given that most experiments have examined gesture's role in L2 acquisition in adults (but see Andrä et al., 2020; Cabrera & Martínez, 2001; Macedonia et al., 2014; Mavilidi et al., 2015; Rowe et al., 2013), the extent to which their findings apply to children is unclear, particularly given that cognitive and language development in childhood affects gesture comprehension and production (Mayberry & Nicoladis, 2000; Volterra et al., 2005). Thus, these experiments should be replicated with children to determine whether the results differ in this population, and if so, how. By addressing these and many other important issues, experimentation can serve as a powerful methodological tool to illuminate gesture's role in L2 acquisition.

Recommended Readings

Jegerski, J., & VanPatten, B. (2013). *Research methods in second language psycholinguistics*. Routledge.

Kelly, S. D., & Lee, A. L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*, *27*(6), 793-807. https://doi.org/10.1080/01690965.2011.581125

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design*. Routledge.

Morett, L. M., & Chang, L. Y. (2015). Emphasising sound and meaning: Pitch gestures enhance Mandarin lexical tone acquisition. *Language, Cognition and Neuroscience*, *30*(3), 347-353. https://doi.org/10.1080/23273798.2014.923105

Nakatsukasa, K. (2019). Gesture-enhanced recasts have limited effects: A case of the regular past tense. *Language Teaching Research*, 1362168819870283. https://doi.org/10.1177/1362168819870283

Stam, G., & Buescher, K. (2018). Gesture research. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 793-809). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59900-1_36

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661-699. https://doi.org/10.1111/j.0023-8333.2005.00320.x

Tellier, M. (2010). The effect of gestures on second language memorisation by young children. In M. Gullberg & K. De Bot (Eds.), *Gestures in language development* (pp. 75-92). John Benjamins Publishing Company. https://doi.org/10.1075/gest.8.2.06tel

Xi, X., Li, P., Baills, F., & Prieto, P. (2020). Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriately mimic target phonetic features. *Journal of Speech, Language, and Hearing Research*, *63*(11), 3571-3585. https://doi.org/10.1044/2020_JSLHR-20-00084

References

Allen, L. Q. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal*, *79*(4), 521–529. https://doi.org/10.1111/j.1540-4781.1995.tb05454.x

Amand, M., & Touhami, Z. (2016). Teaching the pronunciation of sentence final and word boundary stops to French learners of English: Distracted imitation versus audio-visual explanations. *Research in Language*, *14*(4), 377–388.

Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & von Kriegstein, K. (2020). Learning foreign language vocabulary with gestures and pictures enhances vocabulary memory for several months post-learning in eight-year-old school children. *Educational Psychology Review*, *32*, 815–850. https://doi.org/10.1007/s10648-020-09527-z

Baills, F., Suárez-González, N., González-Fuente, S., & Prieto, P. (2019). Observing and producing pitch gestures facilitates the learning of Mandarin Chinese tones and words. *Studies in Second Language Acquisition*, *41*, 33–58.

Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence of L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, *30*(2), 225–251. https://doi.org/10.1017/S0272263108080327

Cabrera, M. P., & Martínez, P. B. (2001). The effects of repetition, comprehension checks, and gestures, on primary school children in an EFL situation. *ELT Journal*, *55*(3), 281–288.

Dahl, T. I., & Ludvigsen, S. (2014). How I see what you're saying: The role of gestures in native and foreign language listening comprehension. *The Modern Language Journal*, *98*(3), 813–833.

Drijvers, L., & Özyürek, A. (2020). Non-native listeners benefit less from gestures and visible speech than native listeners during degraded speech comprehension. *Language and Speech*, *63*(2), 209–220.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test* (4th ed.). Pearson.

Efron, D. (1941). *Gesture and environment: A tentative study of some of the spatio-temporal and linguistic aspects of the gestural behavior of eastern Jews and southern Italians in New York City living under similar as well as different environmental conditions*. King's Crown Press.

Eskildsen, S. W., & Wagner, J. (2013). Recurring and shared gestures in the L2 classroom: Resources for teaching and learning. *European Journal of Applied Linguistics*, *1*(1), 139–161. https://doi.org/10.1515/eujal-2013-0007

Garcia-Gamez, A. B., & Macizo, P. (2019). Learning nouns and verbs in a foreign language: The role of gestures. *Applied Psycholinguistics*, *40*(2), 473–507. https://doi.org/10.1017/S0142716418000656

Gluhareva, D., & Prieto, P. (2017). Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations. *Language Teaching Research*, *21*(5), 609–631. https://doi.org/10.1177/1362168816651463

Graziano, M., & Gullberg, M. (2013). Gesture production and speech fluency in competent speakers and language learners. *Proceedings of Tilburg Gesture Research Meeting (TiGeR) 2013*.

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, *56*(1), 155–196. https://doi.org/10.1111/j.0023-8333.2006.00344.x

Gullberg, M., Roberts, L., & Dimroth, C. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics in Language Teaching*, *50*, 239–276. https://doi.org/10.1515/iral-2012-0010

Hannah, B., Wang, Y., Jongman, A., Sereno, J. A., Cao, J., & Nie, Y. (2017). Cross-modal association between auditory and visuospatial information in Mandarin tone perception in noise by native and non-native perceivers. *Frontiers in Psychology*, *8*, 2051.

Hirata, Y., & Kelly, S. D. (2010). Effects of lips and hands on auditory learning of second-language speech sounds. *Journal of Speech, Language, and Hearing Research*, *53*, 298–310. https://doi.org/10.1044/1092-4388

Hirata, Y., Kelly, S. D., Huang, J., & Manansala, M. (2014). Effects of hand gestures on auditory learning of second-language vowel length contrasts. *Journal of Speech, Language, and Hearing Research*, *57*(6), 2090–2101. https://doi.org/10.1044/2014_JSLHR-S-14-0049

Hoetjes, M., & Van Maastricht, L. (2020). Using Gesture to Facilitate L2 Phoneme Acquisition: The Importance of Gesture and Phoneme Complexity. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.575032

Huang, X., Kim, N., & Christianson, K. (2019). Gesture and vocabulary learning in a second language. *Language Learning*, *69*(1), 177–197.

Hudson, N. (2011). *Teacher gesture in a post-secondary English as a second language classroom: A sociocultural approach*. [Unpublished doctoral dissertation]. University of Nevada, Las Vegas.

Iizuka, T., Nakatsukasa, K., & Braver, A. (2020). The efficacy of gesture on second language pronunciation: An exploratory study of handclapping as a classroom instructional tool. *Language Learning*, *70*(4), 1054–1090.

Jegerski, J., & VanPatten, B. (2013). *Research methods in second language psycholinguistics*. Routledge.

Kelly, S. D., Bailey, A., & Hirata, Y. (2017). Metaphoric gestures facilitate perception of intonation more than length in auditory judgments of non-native phonemic contrasts. *Collabra: Psychology*, *3*(1), 7.

Kelly, S. D., Hirata, Y., Manansala, M., & Huang, J. (2014). Exploring the role of hand gestures in learning novel phoneme contrasts and vocabulary in a second language. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00673

Kelly, S. D., & Lee, A. L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*, *27*, 793–807.

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes*, *24*(2), 313–334. https://doi.org/10.1080/01690960802365567

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kida, T. (2008). Does gesture aid discourse comprehension in the L2? In S. G. McCafferty & G. Stam (Eds.), *Gesture: Second language acquistion and classroom research* (pp. 131-156). Routledge.

Kimura, D., & Kazik, N. (2017). Learning in-progress: On the role of gesture in microgenetic development of L2 grammar. *Gesture*, *16*(1), 127–151.

Krönke, K.-M., Mueller, K., Friederici, A. D., & Obrig, H. (2013). Learning by doing? The effect of gestures on implicit retrieval of newly acquired words. *Cortex*, *49*, 2553–2568. https://doi.org/10.1016/j.cortex.2012.11.016

Kushch, O., Igualada, A., & Prieto, P. (2018). Prominence in speech and gesture favour second language novel word learning. *Language, Cognition and Neuroscience*, *33*(8), 992–1004. https://doi.org/10.1080/23273798.2018.1435894

Laurent, A., Nicoladis, E., & Mayo, M. del P. G. (2015). Gesture restriction affects French-English bilinguals' speech only in French. *Bilingualism*, *18*(2), 340.

Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, *54*(1), 79–117. https://doi.org/10.1111/j.1467-9922.2004.00249.x

Li, P., Baills, F., & Prieto, P. (2020). Observing and producing durational hand gestures facilitates the pronunciation of novel vowel-length contrasts. *Studies in Second Language Acquisition*, *42*(5), 1015–1039.

Llanes-Coromina, J., Prieto, P., & Rohrer, P. (2018). Brief training with rhythmic beat gestures helps L2 pronunciation in a reading aloud task. *Proceedings of the 9th International Conference on Speech Prosody*, 498–502.

Macedonia, M. (2014). Bringing back the body into the mind: Gestures enhance word learning in foreign language. *Frontiers in Psychology*, *5*, 1467.

Macedonia, M., Bergmann, K., & Roithmayr, F. (2014). Imitation of a pedagogical agent's gestures enhances memory for words in second language. *Science Journal of Education*, *2*(5), 162–169.

Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education*, *5*, 196–211. https://doi.org/10.1111/j.1751-228X.2011.01129.x

Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping*, *32*(6), 982–998. https://doi.org/10.1002/hbm.21084

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design*. Routledge.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mavilidi, M.-F., Okely, A. D., Chandler, P., Cliff, D. P., & Paas, F. (2015). Effects of integrated physical exercises and gestures on preschool children's foreign language vocabulary learning. *Educational Psychology Review*, *27*(3), 413–426.

Mayberry, R. I., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Current Directions in Psychological Science*, *9*(6), 192–196.

Mayer, K. M., Yildiz, I. B., Macedonia, M., & von Kriegstein, K. (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology*, *25*(4), 530–535. https://doi.org/10.1016/j.cub.2014.11.068

McCafferty, S. G. (1998). Nonverbal expression and L2 private speech. *Applied Linguistics*, *19*(1), 73–96.

McCafferty, S. G. (2002). Gesture and creating zones of proximal development for second language learning. *The Modern Language Journal*, *86*(2), 192–203. https://doi.org/10.1111/1540-4781.00144

McCafferty, S. G. (2006). *Gesture and the materialization of second language prosody*.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*(3), 350–371. https://doi.org/10.1037/0033-295X.92.3.350

McNeill, D., & Levy, E. (1982). Conceptual representations in language activity and gesture. *Speech, Place, and Action*, *271*, 295.

Morett, L. M. (2014). When hands speak louder than words: The role of gesture in the communication, encoding, and recall of words in a novel second language. *The Modern Language Journal*, *98*(3), 834–853. https://doi.org/10.1111/modl.12125

Morett, L. M. (2018). In hand and in mind: Effects of gesture production and viewing on second language word learning. *Applied Psycholinguistics*, *39*(2), 355–381. https://doi.org/10.1017/S0142716417000388

Morett, L. M. (2019). The Power of an image: Images, not glosses, enhance learning of concrete L2 words in beginning learners. *Journal of Psycholinguistic Research*, *48*, 643–664. https://doi.org/10.1007/s10936-018-9623-2

Morett, L. M., & Chang, L.-Y. (2015). Emphasising sound and meaning: Pitch gestures enhance Mandarin lexical tone acquisition. *Language, Cognition and Neuroscience*, *30*(3), 347–353. https://doi.org/10.1080/23273798.2014.923105

Nakatsukasa, K. (2016). Efficacy of gestures and recasts on the acquisition of locative prepositions. *Studies in Second Language Acquisition*, *38*(4), 771–799. https://doi.org/10.1017/S0272263115000467

Nakatsukasa, K. (2019). Gesture-enhanced recasts have limited effects: A case of the regular past tense. *Language Teaching Research*, 1362168819870283.

Porter, A. (2016). A helping hand with language learning: Teaching French vocabulary with gesture. *The Language Learning Journal*, *44*(2), 236–256.

Repetto, C., Pedroli, E., & Macedonia, M. (2017). Enrichment effects of gestures and pictures on abstract words in a second language. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02136

Roberge, C., Kimura, M., & Kawaguchi, Y. (1996). Nihongo no Hatsuon Shidoo: VT-hoo no Riron to Jissai [Pronunciation training for Japanese: Theory and practice of the VT method]. *Tokyo: Bonjinsha*.

Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology*, *38*, 109–117. https://doi.org/10.1016/j.cedpsych.2012.12.001

Smotrova, T. (2017). Making Pronunciation Visible: Gesture In Teaching Pronunciation. *TESOL Quarterly*, *51*(1), 59–89. https://doi.org/10.1002/tesq.276

Smotrova, T., & Lantolf, J. P. (2013). The function of gesture in lexically focused L2 instructional conversations. *The Modern Language Journal*, *97*, 397–416. https://doi.org/10.1111/j.1540-4781.2013.12008.x

So, W. C., Sim Chen-Hui, C., & Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, *27*(5), 665–681. https://doi.org/10.1080/01690965.2011.573220

Stam, G., & Buescher, K. (2018). Gesture research. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 793-809). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59900-1_36

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*, 661–699. https://doi.org/10.1111/j.0023-8333.2005.00320.x

Sweller, N., Shinooka-Phelan, A., & Austin, E. (2020). The effects of observing and producing gestures on Japanese word learning. *Acta Psychologica*, *207*, 103079.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8*(2), 219–235. https://doi.org/10.1075/gest.8.2.06tel

Tellier, M. (2010). The effect of gestures on second language memorisation by young children. In M. Gullberg & K. De Bot (Eds.), *Gestures in language development* (pp. 75-92). John Benjamins Publishing Company. https://doi.org/10.1075/gest.8.2.06tel

van Compernolle, R. A., & Smotrova, T. (2017). Gesture, meaning, and thinking-for-teaching in unplanned vocabulary explanations. *Classroom Discourse*, *8*(3), 194–213.

Volterra, V., Caselli, M. C., Capirci, O., & Pizzuto, E. (2005). Gesture and the emergence and development of language. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates*, (pp. 3-40). Lawrence Erlbaum Associates, Publishers.

Wagner, E. (2010). The effect of the use of video texts on ESL listening test-taker performance. *Language Testing*, *27*(4), 493–513.

Wang, W., & Loewen, S. (2016). Nonverbal behavior and corrective feedback in nine ESL university-level classrooms. *Language Teaching Research*, *20*(4), 459–478.

Xi, X., Li, P., Baills, F., & Prieto, P. (2020). Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriately mimic target phonetic features. *Journal of Speech, Language, and Hearing Research*, *63*(11), 1–15. https://doi.org/10.1044/2020_JSLHR-20-00084

Yuan, C., González-Fuente, S., Baills, F., & Prieto, P. (2019). Observing pitch gestures favors the learning of Spanish intonation by Mandarin speakers. *Studies in Second Language Acquisition*, *41*(1), 5–32. https://doi.org/10.1017/S0272263117000316

Zhang, Y., Baills, F., & Prieto, P. (2020). Hand-clapping to the rhythm of newly learned words improves L2 pronunciation: Evidence from training Chinese adolescents with French words. *Language Teaching Research*, *24*(5), 666–689. https://doi.org/10.1177/1362168818806531

Zhen, A., Van Hedger, S., Heald, S., Goldin-Meadow, S., & Tian, X. (2019). Manual directional gestures facilitate cross-modal perceptual learning. *Cognition*, *187*, 178–187.

Zheng, A., Hirata, Y., & Kelly, S. D. (2018). Exploring the effects of imitating hand gestures and head nods on L1 and L2 Mandarin tone production. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2179–2195. https://doi.org/10.1044/2018_JSLHR-S-17-0481

Study Box 1

Morett, L. M., & Chang, L. Y. (2015). Emphasising sound and meaning: Pitch gestures enhance Mandarin lexical tone acquisition. *Language, Cognition and Neuroscience, 30* (3)*,* 347-353.

Background

This experiment examined how observing gestures conveying phonological information (lexical tone pitch contours) and semantic information (word meanings) during L2 word learning affected subsequent identification of novel speech sounds and meanings of L2 words. It expanded upon the findings of previous experiments showing that gestures conveying word meanings facilitate learning of phonologically distinct but not phonologically similar L2 words (Kelly & Lee, 2012), and that gestures conveying other non-native speech sounds fail to facilitate their perception (Hirata et al., 2014; Hirata & Kelly, 2010; Kelly et al., 2014).

Research questions (adapted from Morett & Chang, 2015, p. 348)

- Does observing gestures conveying the pitch contours of lexical tones during L2 word learning facilitate identification of them in learned and novel L2 words?
- Does observing gestures conveying referents during L2 word learning facilitate association of L2 words with their meanings?

Method

Participants consisted of 57 young adult monolingual English speakers with no knowledge of Mandarin. Participants first completed a lexical tone identification pretest in which they heard to-be-learned Mandarin words and guessed their lexical tones by pressing buttons labeled with their pitch contours. Participants next completed a brief training in which each of the Mandarin lexical tones was described verbally and presented visually via an image of its pitch contour and aurally via a vowel present in both Mandarin and English. Participants then learned six pairs of Mandarin words differing minimally in lexical tone (e.g., hui1-hui2) representing each possible tone pairing in one of three conditions: pitch gesture, semantic gesture, or no gesture.

Participants assigned to the pitch gesture condition observed a speaker producing gestures conveying the pitch contours of lexical tones in L2 words; participants assigned to the semantic gesture condition observed a speaker producing gestures conveying the meanings of L2 words; and participants assigned to the no gesture condition observed a speaker remaining still. Following a five-minute break, participants completed a posttest in which they heard learned and novel Mandarin words and identified their lexical tones by pressing buttons labeled with their pitch contours and then heard learned Mandarin words and identified their meanings by pressing buttons corresponding to the English gloss of the presented Mandarin word, with the English gloss of the Mandarin word differing minimally from it in lexical tone as a distractor.

Results

Participants in both the pitch and no gesture conditions improved in lexical tone identification from chance-level accuracy at pretest to above chance-level accuracy at posttest, whereas participants assigned to the semantic gesture condition failed to improve in lexical tone identification from pretest to posttest, performing at chance on both. In addition, participants in

the pitch gesture condition identified the English meanings of Mandarin words with above-chance accuracy following learning, exceeding the accuracy of participants in the semantic and no gesture conditions, who identified their meanings with below-chance accuracy following learning. Based on these findings, the authors concluded that gestures conveying non-native phonemes enhance differentiation between the meanings of L2 words differing minimally in them, whereas gestures conveying word meanings hinder discrimination between non-native phonemes in L2 words containing them.

Study Box 2

Xi, X., Li, P., Baills, F., & Prieto, P. (2020). Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriately mimic target phonetic features. *Journal of Speech, Language, and Hearing Research*, *63*(11), 1–15.

Background

This experiment examined how producing gestures conveying phonetic features congruent with non-native phonemes vs. gestures conveying phonetic features incongruent with non-native phonemes affects their perception and production in L2 words containing them. It expands upon the findings of previous experiments showing that observing gestures conveying phonetic features of non-native phonemes enhances their pronunciation in all (Amand & Touhami, 2016) or some cases (Hoetjes & van Maastrict, 2020).

Research questions (adapted from Xi, Li, Baills, & Prieto, 2020, p. 3575)

- Does observing gestures conveying an air burst improve acquisition of non-native aspirated plosives, which have this feature, relative to no gesture?
- How does observing gestures conveying an air burst affect acquisition of non-native aspirated affricates, which lack this feature, relative to no gesture and the effect of gesture on aspirated plosives, which have this feature?

Method

Participants consisted of 50 adult Catalan speakers with no knowledge of Mandarin. Participants first watched a familiarization video in which they were introduced to Mandarin aspirated consonants. Next, they completed a pretest in which they identified whether pairs of Mandarin words differing minimally in aspirated consonants had aspiration or not and repeated Mandarin words with and without aspirated consonants. They then completed a training session in which they viewed videos of a Mandarin speaker producing aspirated consonants either with or without gestures conveying brief air bursts. Finally, they completed a posttest identical to the pretest.

Pronunciation accuracy was determined by having five native Mandarin speakers rate it for the entire word as well as the accuracy of the consonantal feature (aspiration for plosives, duration of frication for affricates). Each of these ratings was averaged across raters.

Results

Perception of Mandarin aspiration consonants was similar for participants receiving training both with and without gestures. In contrast, production of Mandarin aspirated plosives improved from pretest to posttest for participants receiving training with gestures but not for participants receiving training without gestures, whereas production of Mandarin aspirated affricates failed to improve for participants in both groups. The authors concluded that the efficacy of observing hand gestures on the learning of non-native phonemes depends on the appropriateness of the form of those gestures relative to the target phonetic feature such that hand gestures benefit L2 pronunciation only when they appropriately mimic phonetic features.

Study Box 3

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661-699.

Background

This experiment examined how observing gestures and facial cues affect L2 listening comprehension in adult L2 learners with different proficiency levels. It built upon the findings of a previous experiment showing that observing gestures and hearing repetitions and comprehension checks enhanced comprehension of stories in English in elementary school children learning English as a foreign language (Cabrera & Martinez, 2001).

Research questions (adapted from Sueyoshi & Hardison, 2005, p. 668)

- Does observing relevant visual cues such as gestures and lip movements enhance comprehension of lectures in the L2?
- Does L2 proficiency level affect the perception of visual cues such as gestures and lip movements during L2 listening comprehension?

Method

Participants consisted of 42 young adult English as a second language speakers divided into two groups based on English proficiency: low-intermediate and advanced. In groups of 5-8, participants received one of three versions of a video recorded lecture in English on ceramics: AV-gesture-face (audiovisual including gestures and face), AV-face (audiovisual including face but not gestures), or audio only. Participants were prohibited from taking notes during the lecture. Immediately after the lecture, participants completed a multiple-choice test measuring their comprehension of its content.

Results

For participants with low-intermediate English proficiency, English lecture comprehension was highest in the AV-gesture-face condition, indicating that this group benefited from observation of gestures in addition to the face. For participants with high English proficiency, however, English lecture comprehension was highest in the AV-face condition, indicating that this group benefited from observation of the face but not gestures. For both groups, English lecture comprehension was lower in the audio-only condition than in the AV conditions, indicating that they benefited from the presence of audiovisual cues regardless of L2 proficiency. These findings suggest that comprehension of L2 spoken discourse is enhanced by observation of accompanying gestures in listeners with low but not high L2 proficiency.

Question

Because the speaker recorded for the video lecture gestured naturally, this study could not differentiate between the effects of observing different types of gestures. How could an experiment be designed to examine the effects of observing different types of gestures on L2 spoken discourse comprehension?