*Studies in Second Language Acquisition* (2023), **45**, 348–369
doi:10.1017/S0272263122000092

RESEARCH ARTICLE

# The additive use of prosody and morphosyntax in L2 German

Nick Henry*

The University of Texas at Austin, Austin, TX, USA
*Corresponding author. E-mail: nhenry@austin.utexas.edu

(Received 24 June 2021; Revised 19 January 2022; Accepted 7 February 2022)

## Abstract

This study investigates whether the use of prosodic cues during instruction facilitates the processing of German accusative case markers. Two groups of third semester L1 English learners of L2 German completed Processing Instruction (PI) with aural input: Learners in the PI+P group heard sentences that included focused prosodic cues; learners in the PI group heard sentences with monotone prosody. The effects of training were assessed through an offline comprehension task, a written production task, and an online self-paced reading (SPR) task. The results for the offline tasks showed that the groups were similar with respect to their offline comprehension and production. The SPR task showed that both groups used case markers to interpret word order online to some extent; however, only the PI+P group did so in all conditions. These results suggest that prosody does play a role in (morpho)syntactic processing, and that covert activation of prosodic structures can facilitate processing during online reading tasks.

## Introduction

Within the second language (L2) sentence processing literature, it has been widely observed that learners have difficulty processing morphosyntactic forms, and L2 learners often rely on lexical-semantic information to comprehend the input, even after they achieve high proficiency (e.g., Keating, 2009; Marinis et al., 2005). While much research has focused on learners' tendency to favor lexical-semantic over morphosyntactic cues when processing online, comparatively little research has explored the effects of prosody on L2 sentence processing, even though such effects are well-attested in the literature on native (L1) speakers (Steinhauer, 2003). As recent research on Processing Instruction (PI; see VanPatten, 2004b, 2015) suggests that the use of prosody can indeed help L2 learners acquire morphosyntactic forms (Henry, Jackson, et al., 2017), the present study seeks to investigate the role of prosodic cues in online (i.e., real-time) processing.

## Background and Motivation

### *Input Processing and Processing Instruction*

One useful framework for viewing learners' processing strategies and their use of varying types of linguistic information is VanPatten's Input Processing model. The Input Processing model proposes that learners filter input and selectively attend to the most salient cues with the highest communicative value (VanPatten, 2004a). This stems from both the need to get meaning and restraints on cognitive resources (Miyake & Friedman, 1998; see Juffs & Harrington, 2011). Morphosyntactic cues are thus often left unprocessed during comprehension in favor of lexical-semantic or word order cues. One consequence of this filtering is that, as described in the Input Processing model's First Noun Principle (FNP), learners tend to overlook case-marking information and process the first noun in a sentence as the subject or agent (LoCoco, 1987; VanPatten, 1984). Subprinciples of the FNP state that learners may also use context or animacy cues instead of case to understand agent-patient relationships (Jackson, 2007; VanPatten & Houston, 1998).

To illustrate, consider the following examples from German. As seen in (1) and (2), grammatical roles are assigned primarily by case information on the nominative and accusative masculine case markers *der* and *den.* As seen in (2) and (3) these case markers allow the same sentence meaning to be expressed using either subject-verb-object (SVO) or object-verb-subject (OVS) word order:

| | | | | | | |
|---|---|---|---|---|---|---|
| (1) | Der$_{\text{NOM}}$ | Hund | hört | die$_{\text{ACC}}$ | Katze. | (SVO) |
| | The | dog | hears | the | cat. | |
| | "The dog hears the cat." | | | | | |
| (2) | Die$_{\text{NOM}}$ | Katze | hört | den$_{\text{ACC}}$ | Hund. | (SVO) |
| | The | cat | hears | the | dog. | |
| | "The cat hears the dog." | | | | | |
| (3) | Den$_{\text{ACC}}$ | Hund | hört | die$_{\text{NOM}}$ | Katze. | (OVS) |
| | The | dog | hears | the | cat. | |
| | "The cat hears the dog." | | | | | |

However, because feminine, neuter, and plural articles are the same in the nominative and the accusative cases, speakers must actively attend to both word-order cues and case cues. In addition, SVO sentences make up between 80–95% of transitive sentences with full NPs (Kempe & MacWhinney, 1998; Schlesewsky et al., 2000). Learners can therefore rely on a first-noun strategy with a high degree of accuracy, particularly if they also use animacy information to override implausible SVO interpretations (see Jackson, 2007). As a result, learners often do not connect the case markers *der* and *den* to their meaning, delaying acquisition of these morphosyntactic forms (LoCoco, 1987).

The instructional application of Input Processing, PI, seeks to change learners' default processing strategies to promote the acquisition of a targeted form. This is achieved primarily through referential Structured Input (SI) activities, a highly specific type of task-essential activity (Loschkey & Bley-Vroman, 1993), which aims to change learners' processing strategies by manipulating input such that they must use a targeted cue to understand the input and complete a task (see Wong, 2004). In a typical SI activity for German case markers, learners are presented with a mix of SVO and OVS sentences that contain no context or plausibility cues like examples (2) and (3). They must then choose between pictures that correspond to the two possible interpretations

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

of the sentence (i.e., a dog hearing a cat, or a cat hearing the dog). Referential activities (like the one presented in the preceding text) contain unambiguous right and wrong answers; while PI also often includes affective activities, which do not, referential activities have been found to be most effective (Marsden & Chen, 2011). Note, too, that PI focuses exclusively on changing processing behavior through input and does not include any output activities.

Over the last two decades, a vast literature on PI has emerged, and it has been shown to increase learners' comprehension and production of a variety of forms, including those related primarily to verbal morphology (Benati, 2001; Cadierno, 1995) and to grammatical role assignment (VanPatten & Cadierno, 1993; VanPatten & Uludag, 2011). The traditional explanation for PI's positive effects is that it changes learners' processing strategies and pushes them to process grammatical forms. This leads to stronger form-meaning connections within the developing system and an increased ability to use the form in both comprehension and production. While research on PI has made important contributions to research on instructed second language acquisition it is important to emphasize that PI also serves as a methodological tool for investigating how learners process particular aspects of the input. That is, research on PI not only adds to research on the effects of instruction but it also acts as an important validation of the Input Processing model, through which it was developed.

### *The effects of training on processing*

Relatively few studies have used cognitive behavioral measures such as self-paced reading (SPR) or eye-tracking to investigate how PI and related instructional interventions change learners' processing strategies, though these methodologies are becoming more common in PI research. As Lee et al. (2020) and Henry (2022) discuss, these studies address an increasingly large number of research questions, but many have centered on the effects of PI compared to other instructional trainings (Benati, 2020a, 2020b; Chiuchiù & Benati, 2020; Henry, 2022; Issa & Morgan-Short, 2019; McManus & Marsden, 2017, 2018; Wong & Ito, 2018) and on aspects of a PI training, such as the delivery of explicit information (EI) or feedback (Dracos & Henry, 2018; Indrarathne & Kormos, 2017; Issa, 2019; Wong & Ito, 2018). While these studies have utilized a variety of methodological approaches, the research has broadly found that PI increases attention and sensitivity to the target forms (Henry, 2022; Issa & Morgan-Short, 2019), reduces the use of nontarget processing strategies (Dracos & Henry, 2018; Wong & Ito, 2018), and/or increases depth or ease of processing (Chiuchiù & Benati, 2020; Lee et al., 2020). Other studies have found mixed effects or shown that training did not affect online processing (Dracos & Henry, 2021; Ito & Wong, 2019).

One recent study by Henry (2022) investigated whether PI pushes learners to process nominative and accusative case markers in German online. Two groups were instructed on German case markers using either PI or a traditional output-focused instruction (TI) and completed an online SPR task. Results showed that the participants had elevated reading times (RTs) on OVS sentences after receiving PI, but not TI, suggesting that they had processed case markers incrementally as native speakers do (Hemforth et al., 1993; Schlesewsky et al., 2000; Schriefers et al., 1995). However, the effect only occurred when the sentence began with a masculine noun (i.e., unambiguous case marking), and not when the masculine noun came after the verb. Interestingly, results also showed that participants had higher RTs on the noun phrases regardless of the sentence condition. This suggested that PI pushed learners to attend to case

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

markers, though they were not yet able to integrate case-marking information rapidly in all circumstances. In light of these results, Henry suggested that it would be useful to investigate whether the use of prosodic cues in PI could facilitate processing in subsequent online tasks, as recent research has shown that both L1 (Grünloh et al., 2011) and L2 learners (Henry, Hopp et al., 2017; Henry, Jackson et al., 2017; Henry, Jackson, & Hopp, 2020) can use prosodic cues to interpret case markers in German.

## Prosody in L1 and L2

Studies in the L1 acquisition literature have shown that child acquirers exploit links between prosody and syntax to acquire difficult morphosyntactic forms. For example, L1 acquisition research on the Competition Model (see MacWhinney, 2001) has investigated how learners process single cues versus cue coalitions, that is, multiple cues that frequently occur together and point toward the same interpretation. This research has shown that, while German children have difficulty using case markers alone (Dittmar et al., 2008), they can use prosodic cues to help them interpret OVS sentences with unambiguous case information (Grünloh et al., 2011). Research has demonstrated that adult L1 speakers use prosody to resolve syntactic ambiguities (Fodor, 1998; Steinhauer et al., 1999), and to guide grammatical role assignment when morphosyntactic information is absent (Weber et al., 2006). Recent evidence also shows that prosody can be used alongside case information to boost the speed of prediction in German (Henry, Hopp, et al., 2017).

L2 research on the use of prosody and syntax is comparatively less well developed. Early research demonstrated that prosody helps novice learners identify constituents (Wakefield et al., 1974) and order them hierarchically (Langus et al., 2012). Further, prosody may make elements of the input more perceptually salient (Carroll, 2004, 2006). Research in the L2 processing literature has also suggested that prosody helps learners develop more nativelike processing routines (Dekydtspotter et al., 2006; Fernández, 2010). For example, Dekydtspotter et al. (2008) found that fourth-semester learners of French attend to the phonological weight of relative clauses in order to interpret them. The authors concluded that prosody is an "integral part of interlanguage processing" (p. 476), and that the ability to use prosody may be crucial for developing nativelike attachment preferences. Other recent research has found that prosody can help L2 learners make predictions online (Foltz, 2021; Henry et al., 2020). For instance, Henry et al. (2020) found that intermediate-high and advanced L2 learners of German were more likely to use case markers to predict upcoming nouns in a sentence when it included prosodic cues that indicated word order.

To date, only a few studies have investigated the effects of using prosodic information during an instructional training (Henry, Jackson, et al., 2017; Martin & Jackson, 2016). One study, Henry, Jackson, et al. (2017), used PI to investigate whether EI and prosody aid the acquisition of German case markers, the target form in the present study. They found that, when EI was excluded from training, learners were better able to comprehend and produce case cues when they had received training with prosody. They concluded that prosody helps learners identify and attend to morphosyntactic forms, either by increasing the perceptual salience of those forms (e.g., by making phonetically reduced forms like definite articles easier to hear), or by highlighting their communicative purpose. This study thus suggests that PI could push learners to process case markers online more effectively if it includes prosodic cues.

### The Present Study

The aim of the present study is to investigate (a) whether training that includes prosodic cues can push learners to process case markers in German and use them to guide interpretation in real time (i.e., to process case markers incrementally), and (b) whether PI with prosody is more effective than PI that does not utilize prosodic cues.

This study compares the effects of two learner groups who received PI. The first is a group of learners from Henry (2022), whose training did not include prosodic cues (PI). The second group of learners received training that did include prosodic cues (PI+P). The outcomes of training are investigated through offline comprehension and production tasks, as well as an SPR task that evaluates changes to processing behaviors. Thus, the present study sheds light on whether PI changes learners' processing strategies under different training conditions. More importantly, it extends previous research on the use of prosodic cues in L2 processing (Foltz, 2021; Henry et al., 2020) and grammar training (Henry, Jackson, et al., 2017; Martin & Jackson, 2016), by demonstrating whether the presence of prosodic cues in training stimuli has effects on online processing. That is, while Henry et al., (2017) looked only at the outcomes of training on comprehension and production accuracy, this study investigates how learners use case markers moment-by-moment using SPR. Finally, the present research provides insight into the proposal by Dekydtspotter et al. (2006) that the ability to activate and use appropriate prosodic structures is integral to developing nativelike processing routines. The research questions for the present study are as follows:

> **RQ1:** To what extent does a PI training that includes prosodic cues lead to more accurate comprehension and production of accusative case markers in German than PI without prosodic cues?
> **RQ2:** To what extent do learners process German accusative case markers incrementally when comprehending sentences online after training with PI with or without prosodic cues?

## Methodology

### Participants

The participants were drawn from eight intact sections of an intermediate level German course at a large northeastern university in the United States. To determine eligibility for the study, participants completed a language background questionnaire. Each participant included in the final analyses met the following criteria: (a) they were native speakers of English with no advanced knowledge of another language; (b) they demonstrated no knowledge of the target form as determined by a score of 50% or less on OVS items in the pretest's sentence interpretation task (explained in the following section); and (c) they completed all the tasks. The final pool of participants ($N = 53$) was divided randomly into two treatment groups: PI ($n = 25$), and PI with prosody (PI+P) ($n = 28$). The PI group is the same group of participants described in Henry (2022).

To ensure comparability between the groups, the participants completed a written, 30-item multiple-choice language proficiency test (University of Wisconsin Testing and Evaluation, 2006), a working memory task based on Waters and Caplan (1996), and a postexperiment vocabulary test that measured word knowledge and gender assignment for the words used in the SPR task. Descriptive statistics for these measures and responses from the language background questionnaire, including several

**Table 1.** Means for screening and proficiency measures (standard deviations in parentheses)

| Training Group | PI | PI+P |
|---|---|---|
| Variable (Range of Possible Scores) | *Mean (SD)* | *Mean (SD)* |
| Age | 19.36 (2.48) | 19.75 (2.08) |
| Time in German-Speaking Country (in Months) | 0.10 (0.29) | 1.95 (7.18) |
| Years of German Instruction | 4.17 (3.36) | 3.63 (1.78) |
| Years of Instruction in a 3rd Language | 1.91 (2.21) | 1.85 (3.19) |
| Self-Rating: Reading Proficiency (1–10) | 6.02 (1.36) | 5.96 (1.55) |
| Self-Rating: Spelling Proficiency (1–10) | 6.54 (1.51) | 5.79 (1.47) |
| Self-Rating: Writing Proficiency (1–10) | 5.06 (1.28) | 5.36 (1.31) |
| Self-Rating: Speaking Proficiency (1–10) | 5.20 (1.71) | 5.64 (1.34) |
| Self-Rating: Listening Comprehension (1–10) | 5.30 (1.31) | 6.11 (1.69) |
| Working Memory: Set Size (0, 2-6) | 3.68 (0.83) | 3.63 (1.19) |
| Working Memory: Words Remembered (0–89) | 65.36 (11.31) | 63.82 (13.94) |
| Proficiency Task Accuracy (0–30) | 13.24 (5.76) | 13.00 (5.62) |
| Vocabulary Test: Word Knowledge (0–72) | 71.08 (1.87) | 71.14 (1.11) |
| Vocabulary Test: Gender Assignment (0–48) | 45.40 (3.04) | 45.21 (3.00) |

self-rated proficiency measures are presented in Table 1. Statistical analyses showed that the groups were similar in age ($t(51) = .622$, $p = .537$), years of German language instruction ($t(49) = 0.724$, $p = .473$), time spent in a German-speaking country ($t(51) = 1.284$, $p = .205$),[1] overall proficiency ($t(51) = 0.586$, $p = .879$), working memory ($t(51) = 0.438$, $p = .663$), word knowledge ($t(51) = 0.151$, $p = .881$), gender assignment ($t(51) = 0.224$, $p = .824$), and all the self-rated proficiency measures (all $p > .05$).

## Materials

### PI Treatment

A complete record of the training materials is found in the supplemental materials. The PI treatment for both groups consisted of EI, a 50-item referential structured input (SI) activity, and two comparatively shorter affective SI activities that aimed to teach the nominative and accusative case markers *der* and *den* in German.[2] The EI and the referential activity were both presented using the computer program E-Prime (Schneider et al., 2012), while the affective activities were administered using pencil and paper to prevent screen-fatigue among participants.

The training began with the EI, which gave the participants basic information about nominative and masculine case markers and OVS word order. Both groups read information about the communicative purpose of OVS word order. Only the PI+P

[1]The statistical results indicated no differences between the groups, but the PI+P group had a comparatively high mean. This was caused by two participants who were otherwise very similar to the other participants in the study. To ensure that these participants did not affect the study's findings, these participants were removed in a separate analysis. This analysis did not change the results for any of the assessments. These participants were therefore included in the analyses reported in the remainder of the study.

[2]The EI and referential SI activity were the same as those in Henry, Jackson, and Dimidio (2017), and were originally adapted from VanPatten et al. (2013). The affective activities were adapted from Farley (2004). Note that the primary difference between referential and affective activities is that, in referential activities, the target form is task essential, and items have a single correct answer; in affective activities, the target form is used but items allow subjective responses that do not necessarily have a correct answer.

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

group was told about the prosodic cues that accompany OVS word order, but both heard an example of an SVO and OVS sentence using the intonation patterns in their respective trainings.

The referential SI activity for both groups consisted of 38 OVS sentences and 12 SVO distractor items placed in a repeating pattern of three OVS sentences and one SVO sentence so that the distractors were evenly spaced throughout the training. Participants heard the sentences through speakers attached to the computer and were simultaneously presented with two pictures corresponding to the two possible interpretations of the sentences (e.g., a cat hearing a dog, or a dog hearing a cat). Participants then selected the picture that corresponded to the sentence they heard using the computer keyboard. After making their selection, participants saw one-word corrective feedback (i.e., CORRECT, or INCORRECT).

Following the referential activity, participants completed two affective SI activities adapted from Farley (2004). Both were focused on relationships with male persons given that the masculine articles were targeted. In the first activity, participants decided if a series of statements applied to their relationships with a good male friend and a male family member. In the second activity, participants read a list of things a supportive spouse would do and ranked them in terms of importance. There were no correct answers in either activity, but they provided the participants with 26 additional OVS sentences and acted as an input flood.

The only difference between the training for the PI and PI+P groups were the audio recordings in the referential activity. Participants in the PI group heard sentences presented with monotone prosody; the PI+P group heard sentences with focused prosody. The recordings were drawn from the training for the +EI and +EI+P groups in Henry, Jackson, et al. (2017). For the monotone condition, a female native speaker of German was instructed to speak as naturally as possible without emphasizing any of the words in the sentence. For the focused prosody condition, the same speaker imagined that she was responding to a direct question about the subject or the object of the sentence. The prosodic cues in the final stimulus set were evaluated using both the German Tones and Break Indices (GToBI) system (Grice & Baumann, 2002) and a phonetic analysis. These analyses indicated that the OVS sentences with focused prosody carried a high pitch accent with a low leading tone on the first noun phrase (NP1). In SVO sentences with focused prosody, there was no pitch accent on NP1, but the nuclear accent fell on NP2. Thus, the analyses confirmed that the sentences in the focused prosody condition conformed to the pitch contours attested in prior literature (see Braun, 2006; Grünloh et al., 2011; Nespor et al., 2008). The analyses also showed that the sentences in the monotone prosody condition did not have any systematic differences in pitch, duration, or intensity, and none of the sentences contained a high pitch accent on NP1. Thus, these sentences, while spoken naturally, did not contain disambiguating prosodic cues, and sentences in the focused condition were more pragmatically appropriate. The prosodic contours and GToBI ratings for a sample item in each condition are displayed in Figure 1. The results of the phonetic analyses are presented in Table S1 in the supplemental materials.

### *Assessment Measures*

The offline effects of treatment were assessed using a written pretest/posttest that included a sentence interpretation task and a picture description task. The sentence interpretation task consisted of 8 experimental SVO/OVS sentences and 12 distractor sentences followed by a comprehension question in English. The comprehension

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

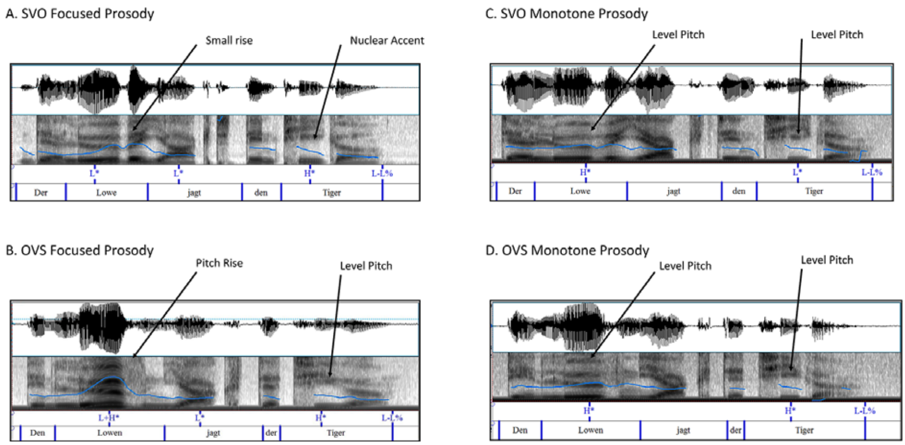


**Figure 1.** Sample waveform and spectrogram with GToBI annotations for training stimuli.

question for the experimental sentences targeted the correct interpretation of grammatical roles as seen in (1):

(1) *Die Oma überrascht der Opa während der Party.*
The$_{ACC}$ grandma surprises the$_{NOM}$ grandpa during the party.
"The grandpa surprises the grandma during the party."
Is the grandpa surprising the grandma? Yes No

Note that the comprehension question is presented in English so that participants could not answer the question by simply matching the case markings from the sentence and the comprehension question.

The picture description task consisted of two target and two distractor picture series. As seen in Figure 2, each series consisted of three pictures, a question prompt, and relevant vocabulary to help participants complete the task.[3] Participants wrote a

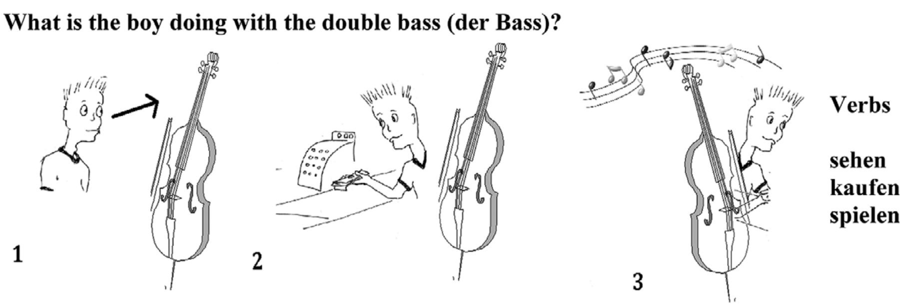


**Figure 2.** Example item from production task in the offline pre-/posttest.

[3]Participants were not limited to the verbs and nouns given to them, but most did use this vocabulary to complete their answer.

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

minimum of one sentence to tell a story in response to the question prompt. In the target picture series, participants saw a main character interacting with a masculine person or object, and thus the sentences specifically elicited use of the masculine articles *der* and *den*.

The online effects of training were assessed using a noncumulative SPR task (Just et al., 1982) using E-Prime (Schneider et al., 2012). The task was administered before and after training and was designed to test sensitivity to case markings through a comparison of RTs on SVO and OVS sentences. As mentioned previously, native speakers tend to display higher RTs at disambiguating regions (i.e., masculine case markers) for OVS sentences when compared with the SVO sentences, indicating incremental use of case marking information. If participants use case marking to assign grammatical roles like native speakers, it is therefore expected that participants would exhibit a similar RT pattern.

During the SPR task, participants first saw a fixation point on the screen. They then pressed the spacebar to begin reading a sentence that was presented phrase by phrase. Participants saw the first phrase in the sentence followed by a series of dashes representing the words in the remainder of the sentence. When participants pressed the spacebar, the first phrase disappeared, and the second phrase appeared. Participants continued in this manner until the end of the sentence. They then answered a Yes/No comprehension question in English to ensure that they attended to the meaning of the sentence.

The SPR task consisted of 72 items: 24 experimental items and 48 filler items. The experimental items were 24 quadruplets containing an NP-V-NP sequence followed by one or two prepositional phrases. The quadruplets were created by varying word order and the position of the masculine noun, resulting in four sentence conditions. Because case cannot be assigned independently of the noun in German, determiners were presented alongside the noun (i.e., NPs were presented together). Thus, sentences were divided into five to seven segments for presentation (see also Hopp, 2006; Jackson, 2008; Schlesewsky et al., 2000). Segments one and three were the disambiguating noun phrases and were the critical regions for analysis in Masculine-First and Masculine-Second sentences, respectively; segments two and four were analyzed for spillover effects. In the following examples, the slashes represent the division of the sentences, bolded segments are the critical regions, and italicized segments are the spillover regions:

(4a) SVO-Masculine First
**Der Opa** / *überrascht* / die Oma / während / der Party.
The$_{NOM}$ grandpa / surprises / the$_{ACC}$ grandma / during / the party.
"The grandpa surprises the grandma during the party."

(4b) OVS-Masculine First
**Den Opa** / *überrascht* / die Oma / während / der Party.
The$_{ACC}$ grandpa / surprises / the$_{NOM}$ grandma / during / the party.
"The grandma surprises the grandpa during the party."

(4c) SVO-Masculine Second
Die Oma / überrascht / **den Opa** / *während* / der Party.
The$_{NOM}$ grandma / surprises / the$_{ACC}$ grandpa / during / the party.
"The grandma surprises the grandpa during the party."

(4d) OVS-Masculine Second
Die Oma / überrascht / **der Opa** / *während* / der Party.
The$_{ACC}$ grandma / surprises / the$_{NOM}$ grandpa / during / the party.
"The grandpa surprises the grandma during the party."

Half the comprehension questions for the experimental stimuli targeted the interpretation of grammatical roles (as seen in the sentence interpretation task), and half targeted the information in the rest of the sentence.

The experimental sentences were split into four lists, and participants only saw one version of each sentence. These lists were controlled so that, within each list, segments were equal in length in each sentence condition. Rather than controlling for the raw frequency of the nouns and verbs used for the experimental stimuli, the participants were trained on the vocabulary so that they would be familiar with it during the SPR task. This method was chosen because pilot data showed that participants were not particularly familiar with many of the words in the SPR task despite them being mostly low-level, high-frequency words found in the participants' coursework. Results of a postexperiment vocabulary test indicated that this method effectively trained the participants' vocabulary knowledge (see Table 1).

## Procedure

The experiment was conducted in three sessions. Session 1 took place in the participants' regular classrooms and included the language background questionnaire, pretest, and proficiency task. Sessions 2 and 3 took place in a lab on campus. Session 2 included the vocabulary training task focusing on the nouns and verbs in the SPR task (see the supplemental materials). In this task, participants saw and heard each word three times with its (written) English translation and repeated the words aloud. In a testing round, they then saw the word, provided a translation of it, and received feedback on their answer. They were required to answer every question correctly before being allowed to move on. After the vocabulary training, they completed the working memory task and the pretest SPR task. In session 3, participants completed a reduced vocabulary training task, in which participants saw each word only twice. They then completed the PI treatment, the written posttest, the posttest SPR task, and a written test to ensure that the participants had retained the vocabulary.

## Data Scoring

### Sentence Interpretation

For the sentence interpretation task, SVO and OVS items were scored separately and given one point for a correct Yes/No answer, and no points for an incorrect answer. The maximum achievable score was four points, one for each target sentence.

### Written Production

For the picture description task, the percentage of accurate responses for the nominative and accusative was scored separately. The scores were computed by dividing the number of accurate responses by the number of obligatory occasions in each participant's response. An obligatory occasion was defined as a point in the sentence, in which the article or pronoun was required to complete the sentence grammatically. The nominative and accusative articles *der* and *den*—as well as their corresponding pronouns *er* and *ihn*, respectively—were considered correct when they accurately described the pictures with respect to grammatical roles. Because participants were not limited in their responses, several participants did not create any obligatory occasions in their responses (e.g., they named the characters or substituted non-masculine nouns to describe objects). These participants did not receive a score and were treated as missing data in the analyses.

#### *SPR Comprehension*

In the SPR task, responses to the comprehension questions for the experimental sentences were recorded by E-Prime (Schneider et al., 2012), along with RTs for each region. The comprehension questions testing grammatical role assignment were assessed as a proportion of accurate responses separately for OVS and SVO sentences.

#### *SPR Reading Times*

The analysis of RTs only included sentences for which the comprehension question was answered correctly,[4] resulting in a loss of 28.7% of the data. Participants who had an overall comprehension rate less than 60% for the entire task or less than 33% for any of the experimental conditions were excluded from analysis. The final analyses for the SPR task therefore included 23 participants in the PI group and 26 from the PI+P group. After these exclusions, the raw RTs were trimmed to remove outliers. First, any RTs below 200 ms or above 4,000 ms were removed from the data. After this, RTs outside a range of +/−3 standard deviations from each participant's overall mean for the experimental items were discarded. The trimming procedures resulted in a loss of 2.9% of the data. After trimming the data, new mean RTs were calculated by participants and condition for each segment of the sentence.

### *Data Analysis*

The data were not normally distributed for any of the tasks. Therefore, the data could not be analyzed using parametric tests. Thus, analyses were conducted using mixed-effects models, which avoid violations of normality and are robust against homoscedasticity and sphericity. The maximal model structure was attempted first (see Barr et al., 2013) and the random effects structure of the models was then reduced when the maximal model did not converge (following Singmann, 2021). The structure of the final model is noted in the results for each analysis. Significant main effects and interactions were explored using post hoc contrasts from the package *emmeans* (Lenth, 2020) and a Bonferroni adjustment for multiple comparisons.

For the interpretation task and accuracy on the SPR comprehension questions, the analyses presented here focus only on OVS sentences and production of the accusative case.[5] For the picture description task, analyses were conducted on production of the accusative case marker. Analyses were performed using the *glmer()* function in the *lme4* package in R using a logit link binomial error distribution. Models were then passed to

[4] A separate analysis was conducted in which all items were included. The overall pattern of results for that analysis did not differ from the results presented here.

[5] Descriptive results for SVO sentences are presented, though not included in the statistical analyses presented here. As can be seen, all learners were highly accurate with SVO sentences and the nominative case throughout the experiment. A separate analysis using the same GLMM procedure can be found in the supplemental materials. This analysis confirmed that there were no effects of Group or interactions containing the factor Group (all $p > .05$) in any of these measures.

the *Anova()* function in the *car* package using contrast coding and type three sums of squares to compute *p* values for fixed effects. The maximal model included fixed effects of Time (Pretest vs. Posttest) and Group (PI vs. TI), and the Time × Group interaction, with by-participant random slopes and intercepts for Time plus the correlation between slopes and intercepts.

RTs for the SPR task were conducted using linear mixed effects models of log-transformed RTs. Separate analyses were performed for each segment, and for Masculine-First and Masculine-Second sentences, as the critical regions differed between them. Models were fit using the *mixed()* function in the R package *afex*[6] (Singmann et al., 2021). The maximal model included the fixed effects of Time (Pretest vs. Posttest), Group (PI vs. TI), and Word Order (SVO vs. OVS) and the interactions between them. The random effects included by-subject and by-item random intercepts and random slopes for Time, Word Order, and the Time × Word Order interaction plus correlations among slopes and intercepts.

## Results

### Sentence Comprehension

The descriptive statistics for the interpretation task are presented in Table 2. The maximal GLMM model yielded no effect for Group ($\chi^2(1) = 0.49$, $p = .484$), but did show a main effect of Time ($\chi^2(1) = 75.16$, $p < .001$) and a marginally significant Time × Group interaction ($\chi^2(1) = 3.74$, $p = .053$). Follow-up pairwise comparisons indicated that both groups made significant gains from pretest to posttest (both $p < .001$), and that there were no significant differences between the groups either at pretest ($p = .247$) or at posttest ($p = .121$).

### Written Production

The group means and standard deviations for the production task are shown in Table 3. The maximal GLMM model yielded a main effect for Time ($\chi^2(1) = 12.24$, $p < .001$), but not for Group ($\chi^2(1) = 0.04$, $p = .839$) or for the Time × Group interaction ($\chi^2(1) = 0.18$, $p = .668$). Pairwise comparisons indicated that both groups improved from pretest to posttest ($p < .001$).

**Table 2.** Descriptive statistics for sentence interpretation task (maximum score of six)

| | SVO Sentences | | | | OVS Sentences | | | |
|---|---|---|---|---|---|---|---|---|
| | *M(SD)* | *95% CI* | *Mdn* | *IQR* | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI | | | | | | | | |
| Pretest | 3.54 (0.8) | 3.21, 3.87 | 4 | 1 | 0.67 (0.62) | 0.42, 0.93 | 1 | 1 |
| Posttest | 3.68 (0.69) | 3.4, 3.96 | 4 | 0.5 | 3.4 (0.91) | 3.02, 3.78 | 4 | 1 |
| PI+P | | | | | | | | |
| Pretest | 3.39 (0.88) | 3.05, 3.73 | 4 | 1 | 0.89 (0.79) | 0.59, 1.2 | 1 | 1.75 |
| Posttest | 3.68 (0.72) | 3.4, 3.96 | 4 | 0 | 2.96 (1.2) | 2.5, 3.43 | 3 | 1 |

*Note:* IQR = Interquartile Range. Values for the 95% confidence interval are bootstrapped using one thousand samples.

[6]The *afex* package acts as a wrapper for the package *lme4* (Bates et al., 2015).

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

**Table 3.** Descriptive statistics for written production task (ratio of correct to obligatory occasions)

| | Nominative Forms | | | | Accusative Forms | | | |
|---|---|---|---|---|---|---|---|---|
| | *M(SD)* | *95% CI* | *Mdn* | *IQR* | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI | | | | | | | | |
| Pretest | 0.95 (0.12) | 0.9, 1.00 | 1.00 | 0.00 | 0.46 (0.4) | 0.3, 0.63 | 0.5 | 0.88 |
| Posttest | 0.99 (0.05) | 0.97, 1.01 | 1.00 | 0.00 | 0.85 (0.32) | 0.72, 0.98 | 1.00 | 0.00 |
| PI+P | | | | | | | | |
| Pretest | 0.96 (0.19) | 0.88, 1.03 | 1.00 | 0.00 | 0.43 (0.46) | 0.25, 0.6 | 0.13 | 1.00 |
| Posttest | 1.00 (0.00) | 1.00, 1.00 | 1.00 | 0.00 | 0.85 (0.31) | 0.73, 0.97 | 1.00 | 0.13 |

*Note:* Values for the 95% confidence interval are bootstrapped using one thousand samples.

**Table 4.** Descriptive statistics for SPR comprehension questions (percentage of correct answers)

| Time | Pretest | | | | Posttest | | | |
|---|---|---|---|---|---|---|---|---|
| | *M(SD)* | *95% CI* | *Mdn* | *IQR* | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI | | | | | | | | |
| Total | 72.99 (4.73) | 70.83, 75.14 | 73.10 | 6.70 | 76 (5.11) | 73.67, 78.32 | 76.90 | 6.65 |
| SVO | 0.78 (0.21) | 0.68, 0.87 | 0.83 | 0.42 | 0.8 (0.15) | 0.74, 0.87 | 0.83 | 0.25 |
| OVS | 0.26 (0.20) | 0.17, 0.35 | 0.17 | 0.25 | 0.44 (0.24) | 0.33, 0.55 | 0.5 | 0.33 |
| PI+P | | | | | | | | |
| Total | 0.74 (0.05) | 0.72, 0.76 | 0.73 | 0.08 | 0.76 (0.05) | 0.74, 0.78 | 0.77 | 0.06 |
| SVO | 0.80 (0.16) | 0.73, 0.86 | 0.83 | 0.33 | 0.76 (0.19) | 0.69, 0.83 | 0.83 | 0.17 |
| OVS | 0.27 (0.23) | 0.18, 0.36 | 0.17 | 0.17 | 0.37 (0.24) | 0.28, 0.46 | 0.33 | 0.33 |

*Note:* Values for the 95% confidence interval are bootstrapped using one thousand samples; values for SVO and OVS sentences reflect only those items in which grammatical role assignment was tested in the comprehension question.

### SPR Comprehension

The means and standard deviations for comprehension accuracy in the SPR task are given in Table 4. The maximal GLMM model yielded a main effect for Time ($\chi^2(1) = 12.15$, $p < .001$). There was no effect for Group ($\chi^2(1) = 0.65$, $p = .421$) or the Time × Group interaction ($\chi^2(1) = 0.51$, $p = .474$). Pairwise comparisons for Time indicated that the participants improved from pretest to posttest ($p = .003$).

Although the PI group outscored the PI+P group on the posttest, both groups failed to reach 50% accuracy on the OVS items in the posttest. To assess whether individual participants had abandoned a strict subject-first strategy, the number of participants reaching 50% accuracy on OVS sentences was evaluated for each group following Henry (2022). In the PI group, this proportion rose from 6 (24%) on the pretest to 15 (60%) on the posttest. In the PI+P group, this figure rose from 6 (23.1%) to 13 participants (50%).

### SPR Reading Times

The analysis of RTs was conducted separately for Masculine-First and Masculine-Second sentences and for each region. For Masculine-First sentences, the critical and spillover segments were Segments 1 and 2. For Masculine-Second sentences, they were Segments 3 and 4.

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

**Table 5.** Mean reading times (SDs) by group and condition for SPR task, Masculine-First items

| Segment | NP1 | V | NP2 | Prep | Final |
|---|---|---|---|---|---|
| Processing Instruction | | | | | |
| SVO, Pretest | 1062 (535) | 898 (436) | 957 (371) | 554 (181) | 918 (377) |
| OVS, Pretest | 1172 (632) | 930 (508) | 1022 (473) | 540 (210) | 888 (385) |
| SVO, Posttest | 1126 (682) | 825 (402) | 1055 (509) | 517 (156) | 861 (387) |
| OVS, Posttest | 1253 (658) | 872 (399) | 1234 (662) | 500 (110) | 872 (343) |
| Processing Instruction with Prosody | | | | | |
| SVO, Pretest | 998 (286) | 941 (317) | 914 (254) | 541 (147) | 825 (245) |
| OVS, Pretest | 1054 (319) | 984 (313) | 962 (318) | 522 (117) | 935 (316) |
| SVO, Posttest | 959 (289) | 840 (277) | 958 (403) | 528 (129) | 790 (304) |
| OVS, Posttest | 1081 (437) | 781 (209) | 1016 (597) | 532 (127) | 892 (362) |

### *RTs for Masculine First Sentences*

The mean RTs and standard deviations for Masculine-First sentences at each segment are displayed by group and condition in Table 5. A full set of descriptive statistics for RTs and the log-transformed RTs are found in Tables S5 and S6 in the supplemental materials.

For the critical segment, Segment 1 (the first NP), the maximal model did not converge. The final model[7] (Table 6) yielded a significant main effect for Word Order. Pairwise comparisons for the SVO-OVS contrast showed that the OVS sentences had higher RTs than SVO sentences ($p = .040$). The estimated marginal means indicate that this effect was driven by larger differences between SVO and OVS sentences in the posttest than in the pretest.

For the spillover segment, Segment 2 (the verb), the maximal model did not converge. The final model[8] (Table 7) yielded only one significant effect, a main effect for Time, which indicated that participants read Segment 2 faster in the posttest than in the pretest ($p = .020$).

**Table 6.** Model results for Segments 1 and 2, Masculine-First Sentences

| | Segment 1 | | | Segment 2 | | |
|---|---|---|---|---|---|---|
| Effect | *Df* | *F* | *p* | *Df* | *F* | *p* |
| Time | 1, 44.11 | 0.08 | .783 | 1, 33.17 | 5.94 | .020* |
| Word Order | 1, 19.76 | 4.83 | .040* | 1, 723.66 | 0.09 | .759 |
| Group | 1, 45.36 | 0.30 | .585 | 1, 43.87 | 0.33 | .569 |
| Time × Word Order | 1, 45.04 | 1.53 | .223 | 1, 711.85 | 0.15 | .696 |
| Time × Group | 1, 44.01 | 0.71 | .403 | 1, 47.23 | 0.23 | .631 |
| Word Order × Group | 1, 46.04 | 0.54 | .465 | 1, 708.76 | 0.25 | .615 |
| Time × Word Order × Group | 1, 44.14 | 0.26 | .613 | 1, 702.36 | 1.33 | .250 |

[7]The final model was: RT.log~Session * WO * Group + (Session * WO || Subject) + (WO | Item). This includes the full fixed-effects structure with by-subject and by-item random effects. “Double bar notation” ( i.e., || ) indicates that random effects did not include the correlation between random intercepts and slopes. See Singmann and Kellen (2019) for a guide on reading mixed model notation.

[8]The final model was: RT.log ~ Session * WO * Group + (Session | Subject) + (Session || Item).

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

**Table 7.** Estimated marginal means for Word Order by Time on Region 1 in Masculine First Sentences

| Word Order | Time | EM Mean | SE | *df* | Lower CL | Upper CL |
|---|---|---|---|---|---|---|
| SVO | Pretest | 2.96 | 0.03 | 89.22 | 2.91 | 3.01 |
| OVS | Pretest | 2.99 | 0.03 | 90.56 | 2.93 | 3.05 |
| SVO | Posttest | 2.95 | 0.03 | 90.30 | 2.90 | 3.01 |
| OVS | Posttest | 3.01 | 0.03 | 85.37 | 2.95 | 3.07 |

### *RTs for Masculine Second Sentences*

The mean RTs and standard deviations for Masculine-Second sentences at each segment are displayed by group and condition in Table 8. A full set of descriptive statistics for RTs and the log-transformed RTs are found in Tables S7 and S8 in the supplemental materials.

For the critical segment, Segment 3 (the second NP), the maximal model did not converge. The final model[9] (Table 9) yielded no significant effects.

For the spillover segment, Segment 4 (the preposition), the maximal model did not converge. The final model[10] (Table 10), yielded a Word Order × Group interaction. Follow-up pairwise comparisons indicated that participants in the PI+P group had higher RTs in OVS sentences than SVO sentences, whereas the PI group did not

**Table 8.** Mean reading times (SDs) by group and condition for SPR task, Masculine-Second items

| Segment | NP1 | V | NP2 | Prep | Final |
|---|---|---|---|---|---|
| Processing Instruction | | | | | |
| SVO, Pretest | 1153 (484) | 988 (552) | 959 (412) | 541 (171) | 895 (333) |
| OVS, Pretest | 1220 (588) | 977 (592) | 922 (502) | 522 (166) | 932 (404) |
| SVO, Posttest | 1117 (584) | 813 (403) | 1028 (484) | 548 (199) | 790 (297) |
| OVS, Posttest | 1103 (534) | 778 (421) | 968 (670) | 511 (189) | 901 (446) |
| Processing Instruction with Prosody | | | | | |
| SVO, Pretest | 1194 (398) | 938 (379) | 986 (351) | 510 (104) | 912 (265) |
| OVS, Pretest | 1100 (415) | 968 (469) | 881 (297) | 555 (131) | 910 (387) |
| SVO, Posttest | 990 (217) | 819 (289) | 998 (465) | 503 (127) | 860 (332) |
| OVS, Posttest | 1065 (443) | 856 (410) | 907 (432) | 682 (569) | 947 (453) |

**Table 9.** Model results for Segments 3 and 4, Masculine-Second Sentences

| | Segment 3 | | | Segment 4 | | |
|---|---|---|---|---|---|---|
| Effect | *Df* | *F* | *p* | *Df* | *F* | *p* |
| Time | 1, 39.02 | 0.01 | .929 | 1, 47.68 | 0.04 | .841 |
| Word Order | 1, 18.40 | 2.67 | .120 | 1, 12.89 | 1.85 | .197 |
| Group | 1, 45.95 | 0.06 | .801 | 1, 44.10 | 0.35 | .558 |
| Time × Word Order | 1, 47.07 | 0.20 | .655 | 1, 699.53 | 0.23 | .634 |
| Time × Group | 1, 45.65 | 0.00 | .980 | 1, 47.70 | 0.03 | .858 |
| Word Order × Group | 1, 79.84 | 0.01 | .926 | 1, 694.38 | 9.83 | .002** |
| Time × Word Order × Group | 1, 45.48 | 0.19 | .669 | 1, 697.83 | 1.28 | .258 |

[9]The final model was: RT.log ~ Session * WO * Group + (Session * WO | Subject) + (Session + WO || Item).

[10]The final model was: RT.log ~ Session * WO * Group + (Session | Subject) + (Session || Item).

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

**Table 10.** Estimated Marginal Means for Segment 4, Masculine-Second Sentences

| Word Order | Time | EM Mean | SE | *df* | Lower CL | Upper CL |
|---|---|---|---|---|---|---|
| Processing Instruction | | | | | | |
| SVO | Pre | 2.70 | 0.02 | 73.76 | 2.65 | 2.74 |
| OVS | Pre | 2.69 | 0.03 | 89.77 | 2.63 | 2.74 |
| SVO | Post | 2.70 | 0.03 | 72.32 | 2.65 | 2.75 |
| OVS | Post | 2.68 | 0.03 | 82.77 | 2.62 | 2.73 |
| Processing Instruction with Prosody | | | | | | |
| SVO | Pre | 2.69 | 0.02 | 74.16 | 2.65 | 2.73 |
| OVS | Pre | 2.72 | 0.02 | 77.61 | 2.67 | 2.77 |
| SVO | Post | 2.67 | 0.02 | 69.74 | 2.63 | 2.72 |
| OVS | Post | 2.74 | 0.03 | 81.29 | 2.68 | 2.79 |

($p = .005$). Estimated marginal means (Table 9) suggest that this effect was driven by the greater differences between SVO and OVS sentences for the PI+P group on the posttest.

## Discussion

To explore whether prosodic cues influence how learners process case markers in L2 German, the present study compared the effects of a traditional PI training (PI) to PI with prosodic cues (PI+P). The assessment measures included offline comprehension and production tasks along with an SPR task to measure changes in online processing.

### *Offline Comprehension and Production*

The offline comprehension and production tasks showed that both groups improved their comprehension accuracy for OVS sentences and the production of the accusative case marker *den*. These results support previous findings in the literature that PI helps learners develop form-meaning connections (e.g., Benati, 2001; VanPatten & Cadierno, 1993), resulting in knowledge that is useful for both comprehension and production.[11] Just as importantly, the PI and PI+P groups improved to a similar degree on offline measures, replicating results from Henry, Jackson, et al. (2017), who found that, when learners received explicit instruction, as they did in this study, prosody did not impact learner performance on offline comprehension and production tasks. Thus, the present research supports their suggestion that L2 learners do not treat intonation and stress cues like lexical-semantic cues during PI, and that prosodic cues do not block attention to morphosyntactic cues.

### *Online Comprehension*

The SPR task explored how learners comprehended sentences in real time. It should first be noted that participants displayed a very strong first-noun strategy in the pretest, interpreting about 80% of the SVO items correctly, but only 25% of the OVS items.

[11]Note that this study is part of a larger project that included a "traditional instruction" control. This control did not improve their comprehension of OVS sentences after training, suggesting that improvement stemmed from PI and not from other factors (see Henry, 2022).

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

While the results of the comprehension questions in the SPR task showed that neither group reached 50% accuracy on OVS items in this task, both groups did improve accuracy on these items. A separate analysis of individuals showed that, in both groups, more than twice as many participants reached 50% accuracy on the posttest. This suggests that, although learners were largely inaccurate on these questions, training did attenuate their tendency to rely on a strict first-noun strategy. Despite this apparent shift, these results stand in stark contrast to the offline comprehension measure, in which participants interpreted OVS items with 85% and 74% accuracy in the PI and PI+P groups, respectively. The difference in scores likely stems from the increased memory load involved in SPR tasks coupled with the fact that participants could not reread any portion of the sentence and had a reduced capacity to apply explicit knowledge.

With respect to the learners' online comprehension patterns, there are several important results. The analysis of RTs showed that the participants from both groups had elevated RTs on OVS sentences in Masculine-First sentences. Further analysis indicated that this effect was driven by higher RTs in OVS sentences on the posttest, although there was no Word Order × Time interaction. As Henry (2022) discusses, this pattern of results indicates that PI had an important, but somewhat limited effect on learners' processing of case markers. Nonetheless, this provides evidence that participants were better able to identify, extract, and integrate case cues after training, representing a movement toward the nativelike processing pattern.

Despite similarities between the groups' processing of Masculine-First sentences, the two groups' processing of OVS sentences in Masculine-Second sentences diverged: Results showed that only the PI+P group also had elevated RTs on OVS sentences in Masculine-Second conditions. This effect provides some indication that the PI+P group actively processed case markers throughout the entire sentence rather than only processing the first NP. It is noteworthy that this effect was delayed until the spillover region, which might indicate less automaticity, stemming from a reduction in available cognitive resources as the sentence is processed. Alternatively, it could be that Masculine-Second sentences are processed less automatically because they are harder to process, for example, because learners tend to process the initial feminine or neuter noun as nominative, even though it is ambiguous. Despite the lack of automaticity, however, it seems that the online effects of the PI+P training were more robust than for the PI training, promoting processing of the target form in all conditions.

### *The Role of Prosody in Online Processing*

Given that learners in PI+P group showed more robust effects in the SPR task, it seems that prosody played a facilitative roll in online processing. Critically, although the PI+P group received aural input that contained a coalition of morphosyntactic and prosodic cues during training, they received no aural input in the SPR task. Thus, the facilitative effect observed in the SPR task implies that the PI+P group was able to use the morphosyntactic cues to activate and apply the appropriate prosodic structures covertly (see Féry, 2005; Fodor, 2002). This covert activation allowed learners to use the coalition between prosodic and morphosyntactic structures additively during silent reading, facilitating processing.

These results support emerging evidence that prosodic information supports syntactic processing (Henry et al., 2020), helping learners identify important cues to word order, create form-meaning mappings, and process those forms online. While previous studies, such as Henry et al., (2020) have shown that prosody plays a role in predictive processing among intermediate high and advanced immersed learners of German, the results of this study are noteworthy because (a) they show that prosody can support morphological processing at relatively low proficiency levels, (b) this can be trained in a relatively short period, and (c) training with aural stimuli transfers to the written modality. It is yet unknown how durable these effects are—as Henry, Jackson, et al. (2017) show, one training is likely not enough to effect long-lasting changes—but these results suggest clear implications for the use of PI and PI with prosody in the short term.

The results also lend critical support to proposals in the sentence processing literature that emphasize the importance of activating prosodic structures during L2 sentence processing. Dekydtspotter et al. (2006), for example, argue that the use of nontarget prosodic structures could be one reason that L2 learners have difficulty processing syntactic structures (Clahsen & Felser, 2006a, 2006b). The present study provides some evidence to support this hypothesis and suggests that the ability to connect syntactic structures to the appropriate prosodic representations does indeed play an important role in the integration of morphosyntactic information online. Notably, these findings also demonstrate that the ability to impose the correct prosodic pattern is not only important for structural ambiguity and attachment preferences but also for morphosyntactic features, like case markers, that are involved in structure building processes during real-time L2 processing.

### Limitations and Directions for Future Research

The present study has several limitations that suggest areas for future research. First, this study did not explicitly manipulate the presence or absence of EI as has been done in previous studies on prosody in German (Henry, Jackson et al., 2017) and on online processing in PI (e.g., Ito & Wong, 2018). While this study represents an important first step in testing the role of prosody in PI and online processing, future research could further elaborate its findings by exploring the independent contribution of prosody and EI. Secondly, the use of SPR may be seen as a limitation, in particular because SPR requires a higher cognitive load than other tasks. Thus, it is difficult to know whether the apparent advantage for the PI+P training appears because of, or in spite of, the use of SPR, and whether similar results would be obtained using a different online measure or whether cognitive differences might play a role (see Dracos & Henry, 2021). Finally, further research should investigate whether the effects of prosody can be traced to its participation in a coalition of cues (as in Henry, Hopp, et al., 2017) or rather stem from increased salience of the target form. In that respect, it would be useful to investigate the relationship between prosody and input enhancement, and to what extent aural and visual input enhancement affects online processing (see Indrarathne & Kormos, 2017).

## Conclusions

The present study provides evidence that the use of prosodic cues during training facilitates (morpho)syntactic processing. Thus, it adds to research suggesting that prosodic cues play a significant role in sentence processing, especially when they form a coalition with other cues. To my knowledge, this is the only research that uses online methodologies to show such effects with L2 learners who have recently been trained with prosodic cues. To the extent that prosody has been underresearched in the L2 acquisition research, it has also been largely ignored in the L2 classroom. This study

https://doi.org/10.1017/S0272263122000092 Published online by Cambridge University Press

suggests that the inclusion of prosodic training may not only help learners with fluency and pronunciation but also with the acquisition and processing of (morpho)syntactic structures (see also Henry, Jackson, et al., 2017). Finally, it should be noted that the present study highlights the utility in combining approaches and methods common in psycholinguistics with instructed L2 acquisition research. Through a psycholinguistic investigation of classroom instruction, the study informs both classroom methods and psycholinguistic theory. While this is not the first study to do so, it represents an important and growing part of L2 acquisition research.

**Supplementary Materials.** To view supplementary material for this article, please visit http://doi.org/10.1017/S0272263122000092.

**Acknowledgments.** This research was supported in part by a National Science Foundation Dissertation Improvement Grant (BCS-1252109). Portions of this research were presented at the 2013 Second Language Research Forum in Atlanta, GA. I would like to thank Carrie Jackson for her endless personal and professional support during the completion of this project. In addition, many thanks are owed to Richard Page, Mike Putnam, Giuli Dussias, Bill VanPatten, Abby Massaro, Adam Baker, Melisa Dracos, Ines Martin, Courtney Johnson Fowler, and to members of the Penn State Center for Language Science, all of whom offered invaluable support and feedback on the present study.

**Data Availability Statement.** The experiment in this article earned an Open Materials badge for transparent practices. The materials are available at https://doi.org/10.18738/T8/58HF6T

## References

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.1177/009286150103500418

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research*, *5*, 95–127. https://doi.org/10.1177/136216880100500202

Benati, A. (2020a). An eye-tracking study on the effects of structured input and traditional instruction on the acquisition of English passive forms. *Instructed Second Language Acquisition*, *4*, 158–179.

Benati, A. (2020b). The effects of structured input and traditional instruction on the acquisition of the English causative passive forms: An eye-tracking study measuring accuracy in responses and processing patterns. *Language Teaching Research.* Advance online publication. https://doi.org/10.1177/1362168820928577

Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language and Speech*, *49*, 451–493. http://www.ncbi.nlm.nih.gov/pubmed/17326588

Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *The Modern Language Journal*, *79*, 179–193.

Carroll, S. (2004). Some general and specific comments on input processing and processing instruction. In B. VanPatten (Ed.), *Processing Instruction: Theory, Research, and Commentary* (pp. 293–309). Lawrence Erlbaum.

Carroll, S. (2006). Salience, awareness and SLA. In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (Issue Gasla, pp. 17–24). Cascadilla Proceedings Project.

Chiuchiù, G., & Benati, A. (2020). A self-paced-reading study on the effects of structured input and textual enhancement on the acquisition of the Italian subjunctive of doubt. *Instructed Second Language Acquisition*, *4*, 235–257. https://doi.org/10.1558/isla.40659

Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, *27*, 107–126.

Clahsen, H., & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42. https://doi.org/10.1017/S0142716406060024

Dekydtspotter, L., Donaldson, B., Edomnds, A. C., Liljestrand Fultz, A., & Petrush, R. A. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English-French learners. *Studies in Second Language Acquisition*, *30*, 453–480. http://journals.cambridge.org/abstract_S0272263108080728

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference*, 33–40.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development*, *79*, 1152–1167.

Dracos, M., & Henry, N. (2018). The effects of task-essential training on L2 processing strategies and the development of Spanish verbal morphology. *Foreign Language Annals*, *51*, 344–368. https://doi.org/10.1111/flan.12341

Dracos, M., & Henry, N. (2021). The role of task-essential training and working memory in offline and online morphological processing. *Languages*, *6*, 24.

Farley, A. P. (2004). *Structured input: Grammar istruction for the acquisition oriented classroom*. McGraw Hill.

Fernández, E. M. (2010). Reading aloud in two languages: the interplay of syntax and prosody. In B. VanPatten & J. Jegerski (Eds.), *Research in second language processing and parsing* (pp. 297–320). John Benjamins.

Féry, C. (2005). Laute und leise prosodie. *IDS Jahrbuch*, *41*, 1–20. http://books.google.com/books?hl=en&lr=&id=IZMKuYZDvD8C&oi=fnd&pg=PA164&dq=Laute+und+leise+Prosodie&ots=M3n4-fpAFA&sig=uqZi571DCzAC8f3x883yuFSjyhg

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, *27*, 285–319.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. *Proceedings of NELS*, *32*, 113–132. http://www.llf.cnrs.fr/Gens/Abeille/nels.final.doc

Foltz, A. (2021). Using prosody to predict upcoming referents in the L1 and the L2. *Studies in Second Language Acquisition*, *43*, 753–780. https://doi.org/10.1017/S0272263120000509

Grice, M., & Baumann, S. (2002). Deutsche intonation und GToBI. *Linguistische Berichte*, *191*, 267–298. http://www.coli.uni-saarland.de/publikationen/softcopies/Grice:2002:DIG.pdf

Grünloh, T., Lieven, E., & Tomasello, M. (2011). German children use prosody to identify participant roles in transitive sentences. *Cognitive Linguistics*, *22*, 393–419. https://doi.org/10.1515/cogl.2011.015

Hemforth, B., Konieczny, L., & Strube, G. (1993). Incremental syntax processing and parsing strategies. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 539–545). Lawrence Erlbaum.

Henry, N. (2022). The offline and online effects of Processing Instruction. Manuscript submitted for publication.

Henry, N., Hopp, H., & Jackson, C. N. (2017). Cue additivity in predictive processing of word order in German. *Language, Cognition and Neuroscience*, *32*, 1229–1249. https://doi.org/10.1080/23273798.2017.1327080

Henry, N., Jackson, C. N., & DiMidio, J. (2017). The role of prosody and explicit instruction in Processing Instruction. *Modern Language Journal*, *101*, 1–21.

Henry, N., Jackson, C. N., & Hopp, H. (2020). Cue coalitions and additivity in predictive processing: The interaction between case and prosody in L2 German. *Second Language Research*. Advance online publication. https://doi.org/10.1177/0267658320963151

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions. *Studies in Second Language Acquisition*, *39*, 401–430. https://doi.org/10.1017/S027226311600019X

Issa, B. I. (2019). Examining the relationships among attentional allocation, working memory, and second language development. *The Routledge Handbook of Second Language Research in Classroom Learning*, 464–479. Routledge. https://doi.org/10.4324/9781315165080-32

Issa, B. I., & Morgan-Short, K. (2019). Effects of external and internal attentional manipulations on second language grammar development. *Studies in Second Language Acquisition*, *41*, 389–417. https://doi.org/10.1017/S027226311800013X

Ito, K., & Wong, W. (2019). Processing instruction and the effects of input modality and voice familiarity on the acquisition of the French causative construction. *Studies in Second Language Acquisition*, *41*, 443–468. https://doi.org/10.1017/S0272263118000281

Jackson, C. N. (2007). The use and non-use of semantic information, word order, and case markings during comprehension by L2 learners of German. *The Modern Language Journal*, *91*, 418–432.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*, 137–166. https://doi.org/10.1017/S0261444810000509

Just, M. A, Carpenter, P. A, & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology. General*, *111*, 228–238. http://www.ncbi.nlm.nih.gov/pubmed/6213735

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, *59*, 503–535. https://doi.org/10.1111/j.1467-9922.2009.00516.x

Kempe, V., & MacWhinney, B. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, *20*, 543–587. https://doi.org/10.1017/S0272263198004045

Langus, A., Marchetto, E., Bion, R. A. H., & Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *Journal of Memory and Language*, *66*, 285–306. https://doi.org/10.1016/j.jml.2011.09.004

Lee, J. F., Malovrh, P. A., Doherty, S., & Nichols, A. (2020). A self-paced reading (SPR) study of the effects of processing instruction on the L2 processing of active and passive sentences. *Language Teaching Research.* Advance online publication. https://doi.org/10.1177/1362168820914025

Lenth, R. V. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version 1.5.3*. https://CRAN.R-project.org/package=emmeans

LoCoco, V. (1987). Learner comprehension of oral and written sentences in German and Spanish: The importance of word order. In B. VanPatten, T. Devorak, & J. F. Lee (Eds.), *Foreign language learning: A research perspective* (pp. 119–129). Newbury House.

Loschkey, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: integrating theory and practice* (pp. 123–167). Multilingual Matters.

MacWhinney, B. (2001). The Competition Model: the input, the context, and the brain. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 69–90). Cambridge University Press.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*, 53–78.

Marsden, E., & Chen, H. Y. (2011). The roles of structured input activities in processing instruction and the kinds of knowledge they promote. *Language Learning*, *61*, 1058–1098. https://doi.org/10.1111/j.1467-9922.2011.00661.x

Martin, I. A., & Jackson, C. N. (2016). Pronunciation training facilitates the learning and retention of L2 grammatical structures. *Foreign Language Annals*, *49*, 658–676. https://doi.org/10.1111/flan.12224

McManus, K., & Marsden, E. (2017). L1 explicit instruction can improve L2 online and offline performance. *Studies in Second Language Acquisition*, *39*, 459–492. https://doi.org/10.1017/s027226311600022x

McManus, K., & Marsden, E. (2018). Online and offline effects of L1 practice in L2 grammar learning: A partial replication. *Studies in Second Language Acquisition*, *40*, 459–475. https://doi.org/10.1017/S0272263117000171

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: working memory as language aptitude. In A. S. Healy & L. E. Bourne (Eds.), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention* (pp. 339–364). Lawrence Erlbaum.

Nespor, M., van de Vijver, R., Schraudolf, H., Shulka, M., Cinzia, A., & Donati, C. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio*, *2*, 1–29.

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In B. Hemforth & L. Konieczny (Eds.), *German Sentence Processing* (pp. 65–93). Kluwer.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2012). *E-Prime v 2.0.10*. Psychology Software Tools Inc.

Schriefers, H., Friederici, A. D., & Kühn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, *34*, 499–520. http://www.sciencedirect.com/science/article/pii/S0749596X85710236

Singmann, H. (2021). *Mixed model reanalysis of RT data.* https://cran.r-project.org/web/packages/afex/vignettes/afex_mixed_example.html

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Schachar, M. (2021). *afex: Analysis of Factorial Experiments.* https://cran.r-project.org/package=afex

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. H. Spieler & E. Schumacher (Eds.), *New Methods in Cognitive Psychology* (pp. 4–31). Psychology Press. https://doi.org/10.4324/9780429318405-2

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*, 142–164. https://doi.org/10.1016/S0093-934X(02)00542-4

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191–196. https://doi.org/10.1038/5757

University of Wisconsin Testing and Evaluation. (2006). *German placement test [Assessment instrument].*

VanPatten, B. (1984). Learners' comprehension of clitic pronouns: More evidence for a word-order strategy. *Hispanic Linguistics*, *1*, 57–67.

VanPatten, B. (2004a). Input Processing in second language acquisition. In B. VanPatten (Ed.), *Processing Instruction: Theory, Research, and Commentary* (pp. 5–31). Lawrence Erlbaum.

VanPatten, B. (Ed.). (2004b). *Processing Instruction: Theory, Research, and Commentary*. Lawrence Erlbaum.

VanPatten, B. (2015). Input processing in adult SLA. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition* (2nd ed., pp. 113–134). Lawrence Erlbaum.

VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, *15*, 225.

VanPatten, B., Collopy, E., Price, J. E., Borst, S., & Qualin, A. (2013). Explicit information, grammatical sensitivity, and the first-noun principle: A cross-linguistic study in processing instruction. *The Modern Language Journal*, *97*, 506–527. https://doi.org/10.1111/j.1540-4781.2013.12007.x

VanPatten, B., & Houston, T. (1998). Contextual effects in processing L2 input sentences. *Spanish Applied Linguistics*, *2*, 53–70.

VanPatten, B., & Uludag, O. (2011). Transfer of training and processing instruction: From input to output. *System*, *39*, 44–53. https://doi.org/10.1016/j.system.2011.01.013

Wakefield, J., Doughtie, E., & Yom, B. (1974). The identification of structural components of an unknown language. *Journal of Psycholinguistic Research*, *3*, 261–269. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:No+Title#0

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology*, *49*, 51–79.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, *99*, 63–72. https://doi.org/10.1016/j.cognition.2005.07.001

Wong, W. (2004). The nature of Processing Instruction. In B. VanPatten (Ed.), *Processing Instruction: Theory, Research, and Commentary* (pp. 33–67). Lawrence Erlbaum.

Wong, W., & Ito, K. (2018). The effects of processing instruction and traditional instruction on L2 online processing of the causative construction in French: An eye-tracking study. *Studies in Second Language Acquisition*, *40*, 241–268. https://doi.org/10.1017/S0272263117000274

**Cite this article:** Henry, N. (2023). The additive use of prosody and morphosyntax in L2 German. *Studies in Second Language Acquisition,* **45,** 348–369. https://doi.org/10.1017/S0272263122000092