Bartın University Journal of Faculty of Education
BUEFAD Volume 14, Issue 2
606-626
dergipark.org.tr/buefad
DOI: 10.14686/buefad.1463760

# Managing Peer-Feedback Resistance in Pre-Service Language Testing and Evaluation Course

***Müberrem Berna Baydar*** [a*] ***& Çiler Hatipoğlu*** [b]

a PhD Student, Middle East Technical University, https://orcid.org/0000-0001-8377-3137 *berna.baydar@metu.edu.tr

b Prof. Dr., Middle East Technical University, https://orcid.org/0000-0002-7171-1673, ciler@metu.edu.tr

**Research Article**
Received: 02.04.2024
Revised: 18.09.2024
Accepted: 15.10.2024

## Abstract

This study examines the interactional practices of senior-year preservice teachers who are tasked with designing and constructing language tests as part of their coursework in the Language Testing and Assessment course, an essential component of the English language teaching curriculum. Building on the task of test construction, preservice teachers engage in group-based, reciprocal peer feedback sessions to identify potential issues in their test items and improve their overall quality, validity, and reliability. To closely examine the interactional dynamics unfolding during these item-review sessions, the study adopts a micro-analytic approach, conversation analysis, exposing the emotionally laden nature of peer feedback. The findings reveal that emotionally charged moments often gave rise to heightened resistance, rejection, and disagreement among peer feedback groups as a response to detailed peer reviews. Accordingly, the emergence of resistance and interactional misalignment among peers led to sequence expansion, involving repeated cycles of evaluation, justification, and recommendation until a joint consensus is achieved. Despite extensive research on resistance encounters in medical interviews, couples counseling, and service exchanges, the field remains notably deficient in studies examining peer feedback practices in the higher education context. Therefore, this study aims to fill this gap by exploring the dynamics of peer feedback interactions among preservice teachers, shedding light on how resistance unfolds and is managed through the emergent phenomenon, *reference to testing principles*.

**Keywords**: peer feedback, pre-service teacher education, resistance, language testing and assessment, conversation analysis.

# Öğretmen Adaylarına Yönelik Ölçme ve Değerlendirme Dersinde Akran Geribildirimine Karşı Direncin Yönetimi

## Öz

Bu çalışma, İngilizce öğretmenliği müfredatının temel bir bileşeni olan ve son sınıf öğretmen adaylarının dersin geçme kriterlerden biri olarak yabancı dil sınavları oluşturdukları yabancı dilde ölçme ve değerlendirme dersi içinde yer almaktadır. Buna dayanarak, öğretmen adayları ayrıca test maddelerinin kalitesini, geçerliliğini ve güvenilirliğini artırmak amacıyla hazırlanan yabancı dil sınavlarındaki olası sorunları belirlemek için gruplar halinde karşılıklı akran geri bildirimi etkileşimlerine dahil olurlar. Diyalojik öğrenme ortamlarında anlamlı etkileşimin merkezi rolü göz önüne alındığında, bu çalışma, test maddelerini inceleme ve değerlendirme oturumları sırasında öğretmen adaylarının etkileşimlerini yakından incelemek için mikro-analitik bir yaklaşım olan söylem çözümlemesinden faydalanmaktadır. Bu oturumlar yansıtıcı diyalog ve katılımı teşvik etmeyi amaçlasa da yapılan analiz, öğretmen adaylarının detaylı akran değerlendirmelerine yüksek direnç ve savunmacılıkla karşılık vermesi nedeniyle akran geribildiriminin duygusal yüklü doğasını ortaya çıkarmıştır. Buna göre, akranlar arasında direnç ve etkileşimsel uyumsuzluğun ortaya çıkması, ortak bir fikir birliğine varılana kadar *değerlendirme, gerekçelendirme ve tavsiyeden* oluşan bir dizi genişlemesine yol açmıştır. Tıbbi görüşmelerde, çift danışmanlığında, geri bildirim sonrası oturumlarda ve hizmet karşılaşmalarında direnç karşılaşmaları üzerine yapılan kapsamlı araştırmalara rağmen, yükseköğretim bağlamında akran geri bildirimi etkileşimlerini inceleyen çalışmalar alanda hala yeterince yer edinmemiştir. Bu nedenle, bu çalışmanın, öğretmen adayları arasındaki akran geri bildirimi etkileşimlerinin dinamiklerini inceleyerek bu boşluğu doldurmaktır. Çalışma, akranlar arasında geri bildirime karşı direncin nasıl ortaya çıktığını ve bu direncin, bir etkileşimsel kaynak olan *sınav ilkelerine yapılan atıflar* aracılığıyla nasıl yönetildiğini ortaya koymaktadır.

**Anahtar kelimeler**: akran geribildirimi, direnç gösterme, yabancı dilde ölçme ve değerlendirme, hizmet öncesi öğretmen eğitimi, söylem çözümlemesi.

To cite this article in APA Style:

Baydar, B. B. & Hatipoğlu, Ç. (2025). Managing Peer-Feedback Resistance in Pre-Service Language Testing and Evaluation Course. *Bartın University Journal of Faculty of Education, 14*(2), 606-626. https://doi.org/10.14686/buefad.1463760

## INTRODUCTION

In the dynamic landscape of language education, effective assessment is indispensable for guiding instructional practices and shaping learning outcomes (Hattie & Timperley, 2007; Hatipoğlu, 2015). However, traditional perspectives often oversimplify assessment by reducing it to a one-dimensional framework that focuses on mere grading and standardized testing (Pastore & Andrede, 2019). Liu and Carless (2006) contrast this conventional notion of assessment as solely a certification tool (summative) with its broader, multidimensional role in promoting learning (formative), emphasizing its more nuanced nature. This substantive paradigm shift directs educators towards embracing a holistic approach, equipping them with both comprehensive skills and modern practices in designing and implementing sound assessment activities, scoring, and interpreting the results tailored to the specific type and purpose of the assessment. (Looney et al., 2017). Central to this transformation is the growing recognition of peer feedback as a key factor in enhancing learning outcomes in higher education (Hatipoğlu, 2017, 2023), a role that closely aligns with the objectives of formative assessment. This process is often characterized by critically reviewing the performance of a product, such as a project, presentation, or task, based on predetermined criteria (Falchikov & Goldfinch, 2000; Tseng & Tsai, 2010) through reflective dialogue and explanatory talk. A typical peer feedback session consists of producing an initial draft receiving feedback and making revisions to the same work accordingly (Nicol et al., 2014). Throughout these stages, learners familiarize themselves with the criteria and standards while engaging in a two-way feedback interaction, which cultivates their understanding of the subject matter (Orsmond, 2002). Unlike traditional teacher-centered feedback, which is often criticized for being untimely and vague (Topping, 2010), peer feedback provides rich, informative comments and diverse perspectives as learners actively engage with one another's work. This dynamic process requires learners to apply established standards to identify problems and offer solutions to their peers' work, thereby engaging in higher-order thinking skills and enhancing their overall learning experience. In this sense, peer feedback embodies a process of collaborative interaction and mutual negotiation within social constructivist approach (Nicol et al., 2014). Consequently, considerable research has extensively explored peer feedback as a prevalent assessment method in higher education, emphasizing its efficiency in learning (Hattie & Timperley, 2007; McConglugue, 2020), principles in designing and implementation (Evans, 2013; Nilson, 2003; Pekrun, 2005), benefits for students (Falchikov & Boud, 2006; Topping, 2010), and its impact on perceptions and attitudes of both students and educators (Metin, 2010; Sierra, 2015). In light of this perspective, Liu and Carless (2006) highlight the importance of cultivating a positive and supportive atmosphere in the classroom, with respect to the 'relational' aspect of teacher feedback literacy. It is essential to recognize that while the practice prepares learners for delivering and receiving constructive feedback, its delicate nature may elicit different emotional reactions that influence both their engagement with the feedback and learners' long-term academic performance (Bell & Arthur, 2018; Hyland, 2000; Richardson, 2005), especially during the transition to higher education (Beaumont et al., 2011; Robinson et al., 2011). The emotional weight of it might stir tensions by confronting interlocutors with challenges to their perceived self-perception and self-esteem, potentially damaging their ideal image of being 'good' (Bleakney et al., 2019). This could easily result in defensive behavior and disagreements between the parties, which eventually yields resistance that halts the progressivity of the talk (Huma et al., 2023). When coupled with harsh and extensive criticism, this resistance might lead to what Pitt and Norton (2017) described as emotional backwash, a sense of failure, demotivation, and disappointment toward academic studies. Therefore, a substantial body of literature has been published on triggered emotional reactions across different professional settings, including academic writing centers (Park, 2014, 2017; Waring, 2005), counseling sessions (Muntigl, 2013; Vehvilainen, 2001), healthcare meetings (Heritage & Sefi, 1992; Perakyla, 2010), and post-feedback sessions between teacher mentors and trainees (Copland, 2010; Waring, 2017). In this respect, the researchers aimed to gain deep insights into why and how resistance emerges in these institutional contexts and how it is managed among different parties involved. Jennifer et al. (2012) discovered that undergraduate students often seek acceptance and approval from their peers and educators, which makes them especially sensitive to critical feedback and susceptible to erosion of their confidence, experiencing high levels of stress and discouragement. Pekrun et al. (2002) also highlighted that adverse reactions such as disagreements, rejections, and resistance may arise as a means of protecting the face and perceived sense of competence and reputation. Additionally, factors such as context, power dynamics, cultural factors, and individual characteristics were identified as contributors to the emergence of resistance (Liu & Carless, 2007). On the other hand, a significant number of researchers (Heritage & Sefi, 1992; Kinnel & Maynard, 1999; Leyland, 2018; Park, 2014; Pudlinksi, 2012; Vehvilainen, 2001) explored interactional dynamics of resistance from a conversation analytic perspective, focusing on how resistance unfolds

sequentially, and the practices employed by students. Their findings indicate that the emergence of resistance often results in sequence expansions and longer negotiations of meaning between interlocutors as it is usually perceived as a dispreferred response (Pomerantz, 2010; Schegloff, 2007; Vehvilainen, 2009), disrupting the progressivity of the talk with delays and mitigations (Heritage, 1984). In this regard, Park (2014) illustrated a typical formulation of resistance, where yeah initially acknowledges the negative feedback, and followed by a contrastive marker, but, which further extends the negotiation process. Subsequently, the epistemic marker, I think, often signals an account and anticipates an upcoming disagreement. Heritage and Sefi, (1992) also identified a direct and overt formulation of resistance in similar contexts. This formulation typically begins with an explicit rejection (no), followed by assertions of knowledge and competence (I know) and includes the articulation of specific conditions (it doesn't work for me) during feedback interactions. Clayman (2013) also distinguishes between the two types of resistance: overt and covert. While overt resistance openly leads to misalignment, covert resistance manifests more subtly, often through passive behaviors such as withholding a response or diverting the topic. In her influential study on advice giving in academic writing centers, Waring (2005) also identified several practices students use to express their resistance, including invoking authority, asserting personal agendas, making irrational refusals, and minimizing the importance of feedback. In terms of managing resistance, she discovered that L2 writing mentors (2007a, 2007b) formulated their feedback using account-giving sequences and proposed suggestions by going general, invoking larger pedagogical principles, and occasionally creating hypothetical situations to mitigate resistance and potential disputes. Additionally, Waring (2007b) highlighted that account-giving is a prevalent strategy for enhancing the validity of feedback and preempting resistance from feedback receivers. Similarly, Chiu (2008) and Coleman (1998) examined the use of accounts, justifications, and explanations in feedback formulations, underlining their informative role and identifying them as natural 'problem solvers'. From another perspective, Vehvilainen (2001) uses disagreements as a step-wise approach to introducing advice. This method involves initially questioning students' ideas to ensure that the advice is delivered pedagogically and is less likely to encounter further resistance.

Several studies (Leung, 2002; Stokoe et al., 2020) also investigated conflict talks and disputes in everyday conversations, shedding light on how resistance emerges and is managed by interlocutors during mundane talk. Strategies such as submission, dominant party intervention, compromise, and withdrawal have been identified as common practices that bring resistance to a resolution and consensus between parties. Because of its substantial consequences for communication processes, organizational behavior, and interpersonal dynamics, the subject of resistance has attracted a great deal of academic study. The majority of research has traditionally focused on the types and effectiveness of feedback practices (Evans, 2013; Nilson, 2003; Pekrun, 2005), with limited attention given to the interactional dynamics of peer feedback, such as the variables influencing its acceptance or rejection. Only recently have studies begun to dive into the various facets of resistance and negotiation among interlocutors, especially in the context of advice formulations. (Leyland, 2018; Park, 2014a; Vehvilainen, 2009; Waring, 2007a; Waring, 2007b). Therefore, this study aims to contribute to the limited yet significant collection of interaction-based studies on peer feedback resistance and the following consensus-building process. Copland, Ma, and Mann (2009) redefine feedback as a multi-unit turn unfolding in combination with assessment, evaluation, counter statements, suggestions, and advice. Therefore, the study focuses on exploring the emergence of resistance among pre-service teachers by particularly examining instances of it within the multi-layered nature of peer feedback exchanges in an undergraduate language assessment and evaluation course. Addressing peer feedback resistance is crucial, as cultivating a supportive environment where students feel safe and comfortable delivering and receiving constructive feedback is essential for continuous improvement and collaboration in academic and professional settings. In this regard, the above review reveals two current research gaps and aims to answer the following research questions:

1.How do pre-service teachers deliver feedback to their peers, and how does it lead to resistance?

2.What interactional practices and strategies are employed by pre-service language teachers to cope with resistance during peer feedback exchanges?

3.Do pre-service teachers display any uptake of feedback following the resolution of resistance and reaching a consensus?

# METHOD

## Research Design

This study adopts an inductive, micro-analytic methodology, namely Conversation Analysis (CA), to examine human social interactions. The primary data source comes from weekly video recordings taken in a classroom setting over a six-week period. These recordings aim to capture and document both verbal & nonverbal interactions of pre-service language teachers. The feedback provided by pre-service teachers is analyzed from an emic perspective, revealing "reference to the testing principles" (RTP) as a recurring phenomenon. Within the CA framework, the phenomenon manifests in three distinct sequential contexts in peer feedback sessions: *problematization, resistance*, and *suggestion*. However, the current study specifically examines the emergence of peer feedback resistance sequences, which unfold immediately after the initial problematization stage, and investigates how arising resistance is resolved through the employment of RTP among pre-service teachers.

## Participants and Research Context

The current study involves 23 senior-year pre-service language teachers enrolled in the foreign language education department at a state university in Türkiye. As a crucial component of their undergraduate curriculum, pre-service teachers took an intensive English language testing and evaluation course (ELTE) implemented in a flipped classroom model for six consecutive weeks during the summer school of 2020. Throughout the course, pre-service teachers met with their professors twice a week for sessions lasting four to five hours, during which theoretical concepts were well integrated with practical application in foreign language testing and assessment. The primary purpose of the ELTE course was to furnish pre-service teachers with a sound knowledge of language assessment while cultivating their assessment literacy, which is essential for their future role as assessors in their careers. Therefore, the first two weeks of the course were devoted to establishing a solid theoretical basis for language assessment. During this phase, the class covered a wide range of topics, including the fundamentals of testing and assessment, various national and international language exams, the intricacies of test construction, techniques of item writing, and the critical concepts of validity and reliability and as well as commonly practiced standards and criteria to evaluate test quality. As of the third week, pre-service teachers begin the practical component of the course, arriving in the classroom having completed the assigned readings and working on the first drafts of their language exams in groups for elementary-level students covering grammar, vocabulary, reading, listening, speaking, and writing sections. In the final three weeks, pre-service teachers are divided into five groups, and each is tasked with providing in-depth feedback on their peers' language exams, focusing on fundamental concepts such as e.g. practicality, validity, reliability, authenticity, and washback.

Table 1. Duration of Video Recordings and Main Focus of Feedback Sessions

| Sessions | Session Type | Duration |
|---|---|---|
| **Session 4** | Feedback on the grammar section | 186 minutes |
| **Session 5.** | Feedback on the vocabulary section | 128 minutes |
| **Session 6.** | Feedback on the reading section | 125 minutes |
| **Session 7.** | Feedback on the writing sections | 122 minutes |
| **Total** | | 12 hours |

Each pre-service teacher submitted written feedback on their peers' first drafts on the university's online system and then provided comprehensive verbal feedback during the peer feedback sessions held in class, as Table 1 illustrates. During these peer feedback sessions, pre-service language teachers are expected to validate their feedback through assigned course materials, such as books and articles covering foreign language testing and assessment, as required by the course professor. As Liu & Carless (2006) suggest, involving students in determining assessment criteria and quality enhances students' awareness of standards and facilitates active engagement during peer feedback interactions. Boud (2000) also highlights the importance of embedding peer feedback practices within the regular course hours to cultivate a culture where giving and receiving feedback becomes a natural learning process, promoting the successful implementation of peer feedback among learners.

**Data Collection**

Throughout these six consecutive weeks, a total of 16 hours of peer feedback sessions were video recorded. However, the study particularly concentrates on these sessions within the course; therefore, twelve hours of data are orthographically transcribed using Transana, a computer-assisted transcription tool, and then adhered to Jefferson's (2004) transcription system conventions to guarantee the standardization and authenticity of naturally occurring conversations. From the emic perspective of CA, a close examination is carried out through *repetitive watching*, *prolonged engagement,* and *persistent observation* (Davis, 1992) by the researcher since transcriptions offer detailed representations of the data for analysis (Perakyla, 2004). For a potential discovery of a phenomenon -a *recurring pattern* employed by pre-service teachers during their peer feedback interactions- the data is presented and analyzed in a data session, drawing on the observations and expertise of the researchers in the field. While the granularity of the transcriptions and video recordings provides the researcher with a comprehensive understanding of the conversational dynamics and nuances, the data session and contributions of experts, through unmotivated examination, further address the validity and reliability issues surrounding the adopted methodology (Antaki et al., 2008).


Figure 1. *Methodological Steps in Conversation Analysis*

**Data Analysis**

Conversation Analysis (CA), an approach for examining social interactions through talk and bodily behavior, was adopted as this study's analytic method of inquiry. With an emphasis on the social actions attained through conversation, CA first surfaced in the mid-1960s and deviated from popular cognitive and empirical approaches by employing *actual talk* as its central analytical unit. To analyze participants' utterances, linguistic forms, and bodily actions, CA utilizes audio and video recordings of naturally occurring conversations and demonstrates a revolutionary perspective on understanding social behavior. The emergence of CA brought attention to the significance of researching not only conversation but also the underlying motivations behind linguistic decisions, famously encapsulated in the question: '*Why this, in this way, right now?*' highlighting the dynamic interplay between social action and its construction through *talk-in-interaction.* In CA, exchanges of talk are regarded as a medium through which actions are performed (Hoey & Kendrick, 2018), and basic organizational features within these exchanges are identified as *turn-taking, turn-construction, sequence,* and *repair* by the early work of conversation analytic studies (Goffman, 1964; Sidnell, 2013; ten Have 2007). A simple exchange between a toddler and the parent is provided below, where we see a child's simple response by simply uttering, 'no,' as children are fascinated with the word itself to convey their refusal or resistance.

```
Parent: it's bedtime, sweetheart. Let's go.
Child:  NO↑, no, no, no.
Parent: uh huh yes, yes we are going to bed (.) no:w
```

What we mean by these two utterances is that they make up an *adjacency pair (AP)*, a basic structural unit in a conversation. While the parent issues a directive to the toddler as a first-pair part of the AP, the child's response, rejection, makes a second-pair part relevant. This is how we will analyze the structural organization of longer instances of resistance among pre-service language teachers during their peer feedback interactions by employing CA.

**Discovered phenomenon:** *Reference to Testing Principles*

Nonetheless, managing resistance isn't always this straightforward, like the one between a parent and a toddler. On the contrary, it manifests itself in longer stretches of talk and negotiations, where finding common ground might be a potentially challenging endeavor. As Waring (2007a, 2007b) put forward in her research on the resistance between writing mentors and students, different types of accounts are formulated within the feedback to counter and forestall resistance between the parties involved in the feedback session. Our study reveals a similar phenomenon practiced by pre-service teachers in the face of resistance during peer feedback interactions. It has been observed that peer feedback groups strategically employ a comprehensive set of twenty-one principles as a guideline to justify their claims when they problematize a test item and provide an account as evidence. These

principles fulfill various functions in the study as a criterion and an account and originate from various sources covered in the classroom. First, these principles guiding the critical review of language exams come from assigned course readings focused on language testing and evaluation, such *as Language Assessment: Principles and Classroom Practices* by D. Brown, *Writing English Language Tests* by J.B. Hughes, *Testing for Language Teachers* by A. Heaton, etc. These readings, including textbooks and articles, aim to equip pre-service teachers with a comprehensive understanding of established theories, principles, and sound practices in language assessment. Moreover, these principles are reinforced, clarified, and repeatedly resorted to by the professor during the theoretical stage of the flipped lesson, where foundational concepts are introduced and discussed over two weeks. Besides, the professors requested pre-service teachers to base their feedback on solid grounds at the beginning of the course, which resulted in continuous and ongoing integration of these principles during the peer feedback sessions. Below is an example from a pre-service teacher in the second feedback session, where feedback is provided to the stem of a problematic reading question and supported by a principle to validate the comment and provide evidence from a reliable source in the field.

UZE: hocam, correct me if I am wrong once I read Heaton page one hundred four or five, I don't remember but he say:s if you have scanning questions, the questions should come first then the text should come later. Because they are going to lead the students to find the correct answer without you know mhm spending too much time. Therefore, we should change the place of this mm I mean it should be placed above I think

Throughout the data, principles are brought forward in peer feedback group comments and formulated with the modal 'should' as they clearly indicate how test items should be constructed according to established norms and standards in foreign language testing and assessment. In another example, the teacher also invokes another principal by going general (Waring, 2017) during the discussion while advising a group of pre-service teachers in session five.

Remember Hughes hm (0.2) what does he tell us hm (0.2) he tells us that we shouldn't eliminate the options by looking at the questions (0.3) right↓

Suffice it to say that these principles guide the pre-service teachers while aligning the theoretical insights of language assessment with hands-on applications, constructing valid and reliable test items. Pre-service teachers are provided a framework for critically reviewing and evaluating the language exams, ensuring that evidence-based practices and pedagogical expertise inform their feedback.

**Research Ethics**

Given the significance of ethical considerations in research on human participants, ensuring participants' anonymity and consent before their involvement is particularly crucial for CA-based studies as they rely on naturally occurring conversations. Throughout the research, ethical issues were considered by informing participants about the procedures and purpose of the conversation analytic study, and consent forms were obtained from voluntary participants. Participants' anonymity and privacy were ensured by allocating pseudonyms, and the study refrained from including visual elements.

## FINDINGS

Extract 1 is from peer feedback session four, devoted to the grammar sections of the language tests. The extract below shows how RTP emerges during the problematization stage of peer feedback sessions, which receives direct resistance from the item-writer group one. On the first question, peer feedback group four (CAN, SEL, ZEK, GUL) delivered their feedback and initiated a problematization on a test item, the first question, intended for sixth-grade pupils, testing simple present tense in multiple-choice format and received resistance from the item writer group one (SAL, YAS, BAR, and OZG).

Extract 1.

**1 PROF: %first &question%&**

%looks around the classroom%

&CAN raises his hand&

2 CAN: **it has a clue**

3 PROF: **hmm hmm**

4 CAN: **Ali likes riding a bicycle but ↑he: (.) of course er its an -s=**

5 PROF: **=*hm hm***

*nods her head*

6 CAN: **we can er estimate it because by- by looking a-at the first question**

7 PROF: **o::ka:y?**

8 CAN: **°its not a clear (.) question (.)because of that°**

9 PROF: **goo::d (.) so >you say that we have too many likes<**

10 **in the question(.)**

11 CAN: **yeah also (.) but (.) Ali likes (.) has the answer of the question(.)** 12 **[actually**

13 PROF: **[#but he does not#**

#reads the stem of the question and fills the blank space#

14 SEL: **[unintelligible**

15 YAS: **[↑but if you are (unintelligible) ↓students**

16 CAN: **[↑you-you can get the s be- because of Ali ee:r then you can delete**

17 **some eer options (.) by looking at (.) this looking=**

18 YAS: **=if you know simple present of course °will be better ° but if you**

19 **do not you can (0.2) er you can just select Σd also**

20 (0.2) Σ--1---

1:CAN looks at the exam paper in front of him

21 CAN: **it is just simple Σpresent tense=**

Σ--->

22 YAS: **=but if the students ~d-don't~ know the students don't know**

23 **they [can't (0.2)**

24 CAN: **[they should kno:w=]**

**25 YAS: =of course they should kno:w↑ [but-]**

**26 SS: [UH HUH]**

Extract 1 begins with the professor's focalization of the first question in the grammar section of item writer group one and sustaining her gaze on the classroom, looking for a preservice teacher among the peer feedback groups to start the feedback session. CAN from the peer feedback group four raises his hand and bids for a turn (Sahlström, 2002), which immediately receives an embodied acknowledgment from the teacher (hmm hmm) in the next turn. In line 2, CAN starts his feedback by directly problematizing the item under discussion with the statement (it has a clue). His contribution not only highlights a potential issue in the focal item but also indicates a violation of a fundamental principle. According to Haladyna (2004), one key principle of test item construction is that items should not contain or be designed in a way that provides clues to the correct answer for students. CAN continues his problemaziton of the test item by reading aloud the first part of the sentence (Ali likes riding a bicycle but HE: (0.2)) and stops at the blank space, which aims to test simple present tense negative form (does not). He produces the certainty marker (of course) and provides the correct answer (it's an s) to the question. The violation of this principle in the question stem offers a clue that makes it easier for sixth-grade students to find the correct answer, *doesn't* among the four distractors of the multiple-choice question, and comprises the validity of the item. Upon CAN's elaboration, the professor issues an embodied acknowledgment token (hmm hmm) in coordination with a head nod. Using inclusive language (we) to refer to his group members or the entire class, CAN proceeds to problematize the focal item in his next turn and brings an account to his problematization. He problematizes the item's design and stem (Ali likes), which facilitates the estimation of the correct answer by making it easier to solve by eliminating the other options (we can estimate it by looking at the first question). In his contribution, CAN invokes his group members, or other preservice teachers as well, as potential task-takers with inclusive (we), and immediately receives another acknowledgment token from the teacher (okay) and produces a negative assessment by labeling the question as (it's not a clear question because of that). In line 10, the professor nods and issues (good) as an explicit positive assessment (Waring, 2008), indicating that CAN's comment was a preferred contribution. She later demonstrates her candidate understanding by reformulating CAN's previous comment (so you say) regarding the question stem, and she brings clarification to the aforementioned clue (we have too many likes in the question) mentioned at the beginning of CAN's turn in line 2. In the next turn, CAN acknowledges and signals his continuation (also) to specify the clue he mentioned in the question stem as (ali likes) that yields the correct answer in the question, which the professor takes up and completes the blank space (but he does not) in a terminal overlap. In line 16, SEL overlaps with YAS, another member of peer feedback four, in the turn-initial position and hence withdraws her utterance. YAS takes the turn and continues with the contrastive marker (but) to signal an upcoming resistance (Park, 2014) to the problematization previously made by CAN. In line 15, YAS invokes non-present actors (students) to formulate her resistance (Leyland, 2021), yet her utterance cuts off when she overlaps with CAN's upcoming turn in line 16. CAN elaborates on his initial claim about the problematic question design, where the first part of the sentence (ali likes) is connected by the marker (but) which signals a clear contrast and strongly suggests that the second part of the question must be in the negative form (does not). This serves as a clue that leads test takers towards the correct answer easily, even if they are not fully confident or knowledgeable in their understanding of simple present tense (you can get the -s because of ali then you can delete some options). The use of the contrastive marker 'but' simplifies the question by limiting the possible options and making the negative form the obvious choice while offering a less challenging task for students.

Upon this, YAS immediately takes the turn and starts formulating her resistance argument by questioning the assumptions underlying the feedback delivered. She acknowledges CAN'S perspective about the advantages of grasping the grammatical structure in answering the question correctly (if the students know simple present, of course, it will be better), yet she still defends the appropriacy of item one's design in terms of its difficulty and assumes that the item is still functional. She rejects the item's perceived easiness and makes the statement (if they do not know, they can also select option d) to counter the argument made by CAN. Here, YAS questions the assumption made by CAN that all students will have a thorough understanding of simple present tense well enough to answer the question easily. She overtly resists CAN's

feedback by implying that the question is a fair measurement of students' understanding of the target subject. In her formulation, she constructs her resistance from the student's perspective and invokes non-present actors (students) to describe a possible student behavior while taking the test (Leyland, 2021; Yöney, 2021). YAS's resistance aims to diminish and minimize CAN's negative feedback and creates counter arguments while defending the overall appropriacy and validity of the question. In the previous lines, the problematization co-constructed by CAN and the professor is still met with overt resistance from one of the item writers, YAS, leading to a sequence expansion and ongoing misalignment between preservice teachers. In lines 21 and 22, they both discuss the complexity of the structure being tested and its appropriacy for test-takers. CAN reinforces that the grammar structure being tested is straightforward -the simple present tense, one of the concepts taught at an early stage for students as it is the fundamental building block of the grammar. So, he implies that the test item is relatively easy if the students understand the basic rules of the simple present. In the next line, we can understand YAS's possible frustration from her incomplete utterances and repetitions and ongoing persistence on the item's complexity and difficulty appropriate for students (if the students don't know they can't). She suggests that the concept might be simple, yet the question still functions and can be challenging for students lacking enough knowledge and grasp of the grammar rules. In line 23, CAN acknowledges her resistance and further responds to her claim from the perspective of test specifications. He reminds her that the question only tests simple present tense structure, which must be already in the student's epistemic domain (they should know) since it is the only grammar structure covered in class according to the test specifications written by the item-writer group. Doing this, CAN receives a shared laughter token from his peers, and in the upcoming turn, YAS issues a proforma agreement by employing a full modified repeat of CAN's previous comment (of course, they should know) yet issues a contrastive marker (but) to pursue further acceptance for her resistance which extract 2 displays below.

Extract 2.

| | | |
|---|---|---|
| **27** | **YAS:** | **[e:r bişey sorabilir miyim]** |
| | | may I ask something |
| **28** | **S2:** | **[(unintelligible)]** |
| **29** | **YAS:** | **[outline says that] the only topic is e:r simple present tense (.) how long** |
| **30** | | **(.) do you prefer (0.2) er to ask °this question° how do you: how would** |
| **31** | | **you ask this question** |
| **32** | **CAN:** | **er I am just suggesting that Ali likes should be removed because of (0.3)** |
| **33** | | **err: it's a clue=** |
| **34** | **PROF:** | **=*hm hm*** |
| | | *nods her head* |
| **34** | **SEL:** | **you measured but (.) sentence you don't measure simple present tense** |
| **35** | **CAN:** | **yeah** |
| **36** | **SEL:** | **you just measured (.) but** |
| **37** | **CAN:** | **yeap this is (unintelligible) [positive or negative** |
| **38** | **SEL:** | **[do the students- ↑do the students know the meaning of but** |

**39 YAS: the-this is not- this is just not the meaning of but**

**40 PROF: bu:t also after HE:: (.) eer (.) they also measured simple present tense**

**41 [anyway right]**

**42 YAS: yes it is**

**43 PROF: w:hy↑ (0.3)**

What follows is a preface with the question (may I ask something) pausing the ongoing dispute and preceding a challenge formulated to CAN by inviting him to articulate his perspective as an item writer on testing the same structure and opening the floor for a specific suggestion (Park, 2017). During resistance encounters, the party resisting the negative evaluation or advice of his/her interlocutor might issue a question to challenge the other party or seek alignment in solving the problem collectively. YAS formulates her question (the outline says the only topic is simple present er, how would you ask this question) with great difficulty as the micro pauses, hesitation markers, and multiple self-repairs reveal between lines 30-32. Here, YAS brings test specifications and outlines as the guideline for constructing the items and draws attention to the alignment between this guideline and the question under review. While a suggestion is relevant to YAS's question as a second pair part, the focus shifts from defending the design of the current item to articulating a specific alternative, a suggestion still indicating some form of implicit resistance by the item writer. Instead of taking responsibility for constructing an item to test the simple present tense and providing a suggestion as requested by YAS, CAN begins his formulation with mitigation (I am just suggesting that) He steers the conversation back to the core issue: whether to keep the stem (Ali likes) as the stem of the question or delete it. In this regard, CAN offers to remove the first part of the question (Ali likes) as it functions as a clue to the correct answer. Although his response does not entirely address YAS's question about the test item, he reformulates his initial assertion as a suggestion, underscoring his stance on the matter, which might cause YAS to reconsider her own stance. In the following line, YAS sustains her gaze on CAN yet does not provide an acknowledgment or rejection of his suggestion. Clayman (2013) asserts that being silent and withdrawing a response can also be marked as resistance. We also don't see any reaction from the other item writers in the same group with YAS; they remain silent throughout this negotiation. In the next turn, CAN receives an embodied acknowledgment token from the professor, and in line 34, SEL, another peer feedback provider sitting next to CAN, takes the turn by directly addressing YAS with the second-person singular pronoun (you) and starts her own problematization on the same item and seeks further acceptance from YAS. Wang (2007b) suggests resistance can result in reformulations and taking the matter from different perspectives during the negotiation process in the ongoing interaction until it resolves. At the same time, YAS and SAL, another member of the item-writer group, turn and sustain their gaze at SEL. In her turn, SEL shows her alignment with her group member CAN by bringing additional support for the initial problematization of the item by reformulating the initial comment. She states that the question tests (but) structure, not the present tense, which adds another layer to the problematization. This time, SEL challenges the validity of the same item by asserting that something is measured, but it wasn't the grammatical construct intended to be measured. Following that, CAN confirms SEL in an overlap and shows his agreement by explaining that the question tests positive or negative forms of aux forms do and does with the connector (but) in the middle of the stem. In line 43, SEL retakes her turn and requests information by issuing a question if the students, hypothetical sixth graders as test-takers (Leyland, 2021), know the meaning and function of (but). As a common practice in resistance encounters, SEL's question reinforces her initial comment in line 34 and challenges YAS by putting her in a position to either confirm or disconfirm. In the next turn, YAS performs a disconfirming response and directly rejects SEL's initial claim with a counter statement, stating that the question does not aim to test whether the students know the meaning of (but) or not. The discussion between groups evolved around the initial problematic connector (but) in the question stem, which initially served as a clue and made the answer predictable for students but, at the same time, revealed another issue: the item doesn't test what it aimed to test, or does not measure what is taught in the classroom, simple present tense. In line 40, the professor, as the third party, intervenes and reminds peer feedback groups that the item-writers also measured the same structure, simple present tense after the connector (but), which receives an immediate confirmation from YAS (yes, it is). Upon

this, the teacher issues a wh-interrogative (why). In her question, the teacher might signal the problem in the question stem and ask for further clarification or justification from the item writer groups. Extract 3 illustrates the teacher's alignment and opening the floor for suggestions to eradicate the problem in the item.

Extract 3.

**44 CAN: #you told us that e:r (0.2) we shouldn't eliminate er the options by looking**

**45 at the question ↑right**

#sustains his gaze at teacher--->55

**46 PROF: correct**

**47 CAN: thi-this helps students to eliminate% [(...)**

%-----points to the exam paper-----%

**48 PROF: [hmm hmm hmm hmm I knew (.) eer that and we should rephrase in a different**

**49 manner (.) good and specific for you I was just listening to you for a-**

**50 for now o:kay**

**52 YAS: #+oka:y#+**

#item writer GÜL take notes#

+YAS looks at the paper in front of her+

**53 PROF: so what do you suggest feedback groups**

After the professor issues a (why) interrogative to groups and invites elaboration on the design of the question, upon a micro (0.3) delay, CAN, who is the first to problematize the item, sustains his gaze on the professor. He starts his multi-unit turn by referring to a past learning event (Can Daşkın, 2017) (you told us), which directly takes the teacher as the primary source for his upcoming comment. While formulating his statement, he makes a direct reference to the violated testing principle (we should not eliminate the options by looking at the question right) and asks for confirmation. CAN references previous instructions given by the professor, which indicates his attempt to validate the feedback he provided and ensure its alignment with those instructions to construct valid and reliable test items. He highlights these principles as guidelines to be followed while constructing valid test items and also reaffirms his stance through the professor. CAN emphasizes the problematic question design, which helps *elimination of the options* once again and shows how this contradicts the professor's earlier instructions during the lecturing weeks of the course. While he displays his knowledge of assessment built upon the professors' in-class teachings, he also substantiates his initial claim by referencing the testing principle and bringing evidence to the consistency and validity of his feedback. While doing this, he positions the professor as an epistemic authority, leveraging her for validation and credibility. In this case, *resorting to authority as a third-party intervention* during disputes is a common strategy for conflict resolution (Leung, 2002); while peer feedback groups are all in equal positions, the professor is superior in her epistemic knowledge domain, and therefore, the interaction dynamics change easily and becomes a matter of adhering to the authoritative guidelines. While CAN employs RTP + and a direct reference to the professor as a resolution to the ongoing conflict, he demonstrates that the violated testing principle is within his domain of knowledge on testing and assessment, acquired through the ELTE course taught by the professor (you told us) during the initial weeks of the summer school, before pre-service teachers start drafting their achievement exams. This also directly brings the professor into the ongoing discussion and encourages her direct involvement in resolving the disagreements between the peer feedback groups and providing further clarifications on the matter as an expert on

language testing and assessment. In line 48, the professor confirms peer feedback group members CAN and SEL and shows her alignment with their feedback with a positive assessment marker `(good)`. In the same turn, she states that they should rephrase the question in a different form with inclusive language `(we should rephrase it in a different manner)`, which might signal her going general as a strategy used in advice formulations (Waring, 2007b). Later, she turns to the item-writer group and explains that she was listening to them. Her statement might indicate that she deliberately created an interactional space for both groups to deliver their feedback, discuss the test items with one another, and share their insights. At the end of the interaction, YAS produces an acknowledgment token `(okay)` to the teacher's contribution in the previous line while her group members take notes. She remains silent, which indicates her recognition of the professor's authority and accordingly withdrawal of her participation. Yet this necessarily does not reveal the item writers' viewpoint or their acceptance of the feedback (Vehvilainen, 2009), which requires analysis in the last week of the course dedicated to self-feedback sessions for item writer groups to reveal the changes they made to their language tests in light of the feedback they received. The final closure is issued by the professor by inviting preservice teachers to present their suggestions to eradicate the problem in the item, which indicates her acceptance and alignment with Can's and his group members' feedback on the problematic test items `(so what do you suggest)` once again.

## DISCUSSION

Extract one displays a detailed examination of how pre-service language teachers engage in peer feedback sessions to review and provide feedback on the validity and reliability of test items constructed by their peers. The focal point is around CAN's problematizing the first question on the grammar section of the exam, prepared by group one, which triggers a series of interactions and longer stretches of talk among peer feedback groups and the professor. CAN formulates his feedback on his identification of a violated testing principle: eliminating the options by looking at the question stem itself. Nonetheless, initially, CAN assumes the role of a test taker, solves the question, and problematizes the design of the item constructed for sixth graders. He asserts that the stem of the question serves as a clue that facilitates the selection of the correct answer easily. Upon his negative feedback on the item, he elaborates on the problem by providing an account (Waring, 2017) of the nature of the issue and the construction of the test item. His questioning of the validity of the test prompts a critical examination of the focal item among preservice teachers, leading to a long process of meaning negotiation, which is inherent in the nature of peer feedback interactions. However, in his initial attempt, CAN failed to establish a mutual agreement between the peer feedback and the item-writer group. This demonstrates that even though pre-service teachers employ a common practice in feedback sessions, *taking the role of a test taker/student* and *account-giving* to justify their claims and provide solid explanations, the analysis shows that "not every explanation is equally acceptable" (Verkuyten, 2000). CAN's initial negative feedback sparks a debate between CAN and YAS, the item writer, leading to the emergence of resistance. Resistance can take many forms and is only natural in influencing the dynamics of the feedback process differently. In his feedback, CAN focuses on whether the question effectively measures the intended grammatical structure, simple present and argues for the necessity of following and adhering to the testing principles while constructing test items. On the other hand, YAS as the item writer displays overt resistance from the perspective of non-present actors, students (Leyland, 2021) with hypothetical conditional, to emphasize the knowledge students require to answer the question. Here, YAS questions the assumption made by CAN on the easiness of the item and defends the appropriacy of the item design in terms of its difficulty, which she perceives as suitable for the target level. While she directly resists the feedback provided by group four, she tries to minimize it by implying that the question is a fair measurement of the target structure. Not surprisingly, YAS's overt resistance resulted in the delay of the agreement, leading to sequence expansion and a more extended negotiation process among pre-service teachers. Upon receiving CAN's feedback, YAS maintained her role and position as the item writer, responding defensively by formulating counterarguments grounded in her understanding of students' knowledge domain and the exam's intended coverage, thereby asserting the test's content validity. While group four upholds their roles as item writers, the peer feedback group undertakes different positions as test takers or feedback providers according to the nature of the ongoing conversation, where both parties pursue acceptance of their arguments. During this process, unfolding resistance delays the acceptance of the feedback and common consensus between the parties, yet it also becomes constructive because it facilitates a deeper analysis of the feedback and the focal question. Besides, the unfolding resistance prompted the peer feedback group four to revise their arguments and justify their claims with solid grounds to address the challenges made by the item-writer group to minimize the efficiency and validity of their feedback. As a consequence of this,

CAN reformulates his initial feedback and invokes the violated testing principle (RTP) with reference to the teacher, a third party in the interaction *(you taught us that)* as a shared epistemic source and *knowledge authority* (Heritage, 2013; Sidnell, 2015) which brings the resistance into a resolution. CAN strategically invokes the violated testing principle as an interactional resource (we shouldn't eliminate the options by looking at the question) and as a valid account that aligns his arguments with established pedagogical norms. His employment of RTP displays his situated understanding of language testing and assessment (Can, 2020), and he successfully manages the ongoing resistance from the item writer group.

In this regard, Firth (1995) asserts that accounts can also function as problem-solving devices in such environments open to disagreements. However, the extracts above clearly show that accounts are susceptible to negotiation among pre-service teachers. Our analysis reveals that students do not readily accept every feedback, even if it is presented with valid accounts/evidence, and it naturally yields resistance and disagreements in longer stretches of talk. The acceptance of the problematization depends on the kind of evidence provided. In this case, the professor's presence and the testing principles serve both as guidelines during the item review process and problem-solving devices in the face of ongoing resistance. At the same time, CAN's feedback is validated by the professor's acceptance and subsequent agreement, which leads to a shift in focus towards problem-solving, and hence, the suggestion phase as a common practice in feedback interactions becomes the next relevant action (so what do you suggest). The suggestion phase aims to eradicate the problem in the test item and leads to the next phase in the peer feedback discussion, the uptake. The present study has disproved the widespread belief that the feedback receivers' resistance signifies a conflict between the interlocutors and could lessen the effectiveness of the feedback. (Rodby, 2002; Seckendorf, 1986). Instead, it has demonstrated how the resistance can encourage further discussion and negotiation of the target issue, ultimately promoting critical engagement with pedagogical practices and enhancing collaborative learning.

**Uptake of the Feedback**

A bulk of research investigates the problems inherent in feedback, especially concerning its delivery and acceptance (Falchikov & Goldfinch, 2000; Orsmond, 2002; Tseng & Tsai, 2010). Nevertheless, the main issue lies in the reception and understanding of the feedback. If feedback is not delivered in the most comprehensible and clear manner, indicating the areas of weakness or opportunities for improvement in the subsequent work, it cannot be utilized effectively (Cushing et al., 2011). Hence, its potential remains unexplored. As an essential component of the learning process, feedback should bridge the gap between the current performance and the desired one (Sadler, 1989). Topping (1999) illustrates the situation by stating that feedback should be effectively used to produce an improved version of the same work, such as an assignment or project, fostering learning. Therefore, the last session was dedicated to pre-service teachers' presentation of the latest version of their language exams, and it clearly exhibits how peer feedback groups made revisions to their language test in line with the feedback they received to increase the quality and validity of their test items. Below, we present an extract from week six, session eight, where all groups took fifteen minutes to present, and YAS elaborated on the problematic test item criticized for violating testing principles and how the group has dealt with the problem.

YAS: First of all, we will start with our grammar questions because it was the first session. err as you can see here on the previous err previous version err exam of the exam. you told us that I mean our friends told us that our first grammar question was problematic. because the err the stem, *Ali likes* was giving clues to students, and that's why we changed it with something else as you said here er as you suggested. You can see the new item here on the left side.

This last session provides us with the uptake of the feedback and withdrawal of the resistance as YAS presents. As the extract displays, the resistance does not extend outside the classroom and YAS and other item writer groups are provided with enough time to revise and make necessary changes to their work. Evans (2013) underlines the importance of providing students with time and opportunities to make necessary revisions to their assignments and incorporate feedback. Otherwise, students usually postpone necessary revisions, resulting in a lack of meaningful feedback utilization (Vehvilainen, 2009). In this regard, YAS highlights that, as the item writer group, they have made the necessary changes to the problematic test item according to the feedback they received from their peers. This shows us that the item writer group has accepted and utilized the feedback by following general standards and rules of language testing and assessment. A culture of assessment is cultivated through the professor and the pre-service teachers working together towards a goal, preparing quality language exams in quality with standard criteria. This kind of peer feedback is formative, constructive, and, hence, effective for learning because it is

"formulated, delivered, and framed" in a format (Cushing et al., 2011) that invites pre-service teachers to display active engagement and their situated understanding of the subject matter, language testing, and assessment.

## CONCLUSION

This study delves into pre-service L2 teachers' classroom interaction in an undergraduate language assessment and testing course within the ELT program at a state university in Türkiye. While traditional teacher-led feedback interactions typically center on the *transmission of information and provision* of comments on students' work (Chen & Yu, 2019; Vehvilainen, 2009), the current study presents a different dynamic; student-led item review interactions within peer feedback groups for their work-in-progress. Based on the conversation analytic approach applied to 12-hour classroom data, the study revealed a phenomenon termed "reference to testing principles" (RTP) and its central role in three main phases of the peer feedback interaction: These stages follow a pattern that begins with *problematizing the test item,* proceeds to *resistance against the problematization* and concludes with *suggesting a solution*, addressing the first research question of the study. This process aligns with the findings of earlier research on feedback interactions (Pudlinksi, 2012; Vehvilainen, 1999; Wang, 2007b), which shows that feedback talk typically consists of multi-unit sequences combined with assessments, suggestions, and counterstatements (Park, 2014a; Vehvilainen, 2009; Waring, 2007a; Waring, 2007b), each performing different functions in the ongoing interaction. Additionally, our data revealed a further function of these multi-unit feedback formulations, especially when combined with the discovered RTP phenomenon: managing and mitigating resistance. When preservice teachers embrace their new role as item reviewers and deliver detailed feedback to item writer groups, they naturally encounter defensive behaviors and resistance, which reflect an epistemic asymmetry between the item writers and the feedback groups (Butler et al., 2010). Therefore, the second research question investigates how resistance emerges and is managed between pre-service teachers, with the findings revealing another pattern consisting of i. initial feedback provided by the peer feedback groups, ii. the emergence of resistance from the item-writers iii. The reformulation of feedback through RTP, and iv. the achievement of mutual consensus. Following the initial feedback sequence, our findings align with existing research (Badem-Korkmaz et al.,2024; Pudlinski, 2012; Vehvilainen, 2009; Waring, 2007b) regarding the practices employed by feedback recipients to signal their misalignment and resistance. In the beginning, pre-service teachers who constructed the test items have explicitly defended their exams by *maintaining their positions* as item writers and reported *specific conditions of their language exam* `(but the outline says)`, *questioning the feedback and peer groups* `(how would you ask this question)`, *minimizing the import of the feedback,* and *making counterclaims.* Building on previous research, item writer groups also *invoked non-present actors* and stated *congruence with L2 learners' level* `(if the students know, the only topic is)` to challenge the peer feedback groups and defend the appropriateness of their test item under review (Pudlinski, 2012). In response to this resistance, peer feedback groups revised their initial feedback and reformulated it with explicit accounts (Waring, 2005, 2007a) tailored to the understanding of the item writers through the use of RTP. In this regard, our study illustrates that accounts, when combined with (RTP) serve not only as justifications (Waring, 2007b) but also as problem-solving devices for resolving resistance. In cases of persistent misalignment, peer feedback groups further integrated their accounts with *third-party interventions* (Leung, 2002; Waring, 2005), where the professor who served as the *knowledge authority* was brought into the interaction to terminate the ongoing disagreement between peer feedback groups.

Considering the face issues among peers and the perception of feedback as merely critical rather than constructive, the emergence of resistance provided ample opportunities for preservice teachers to engage in explanatory, dialogic feedback talk and negotiate their current understanding of language testing and assessment. In this respect, the current study aligns with previous research (Cowan & Crème, 2005; Topping & Ehyl, 1998; Yang, 2015), which states that peer-to-peer feedback interactions are multi-faceted processes that evolve through negotiations, discussions, and co-construction of rules.

Through these dialogic discussions centered on pre-service teachers' language tests, the study emphasizes the formative nature of peer feedback. In preparing language exams and providing feedback, they collaborate by sharing insights and strategies and display agency by actively engaging in the review process and making decisions about the test items. This approach enriches pedagogy by fostering autonomy, agency, and collaboration among preservice teachers. They are surrounded with ample opportunities to improve their feedback-giving skills and assessment abilities by pinpointing and addressing potential issues in the test items, thereby enhancing the overall quality and efficiency of assessment practices. This process also contributes to their feedback literacy which involves *appreciating the feedback, making judgments, managing emotions* and *taking action based on the*

*feedback* (Boud & Carless, 2018). Building on these improvements, the study further investigates how preservice teachers' refined feedback-giving skills and assessment abilities contribute to effective use of feedback. To address the third research question on the uptake of the feedback, we have discovered that pre-service teachers reached mutual consensus through this reflective practice, and the resistance did not extend outside the classroom but was withdrawn, as illustrated in the last week of the course dedicated to the presentation session. The study also addressed the traditional criticism of feedback as nonspecific, vague, and untimely for further utilization (Crook et al., 2012; McConlugue, 2020). This critique is addressed in our findings, where the item writers presented the final versions of their language exams and explained how they utilized the feedback they received by making necessary revisions to the problematic test items. This reflective process guided by foundamental principles in language testing and assessment illustrates their emerging feedback literacy, particularly their ability to justify how feedback is used, (including which comments are accepted, adapted or rejected) and to exercise self-evaluative judgment when making informed decisions. This was made possible through their engagement with the assessment criteria and standards, which they had been explicitly trained to interpret and internalize during the initial weeks of the course dedicated to the theoretical foundations of language testing and assessment. Students' familiarization with these standards enabled them to translate abstract criteria into guiding principles for revising their work, an essential condition for fostering their *proactive agency* in engaging with the feedback. To summarize, through a close examination of pre-service teachers' peer feedback interactions, we have presented a case where an interactional resource, reference to testing principles, is used strategically by learners to inform their feedback practices, showing their situated understanding of effective test design and reformulate their feedback to manage any arising disputes and resistance. By collectively establishing guidelines based on shared epistemic resources, pre-service teachers cement the understanding of the nature and severity of the problems on language exams (Can, 2020), resolve any rising conflicts and disagreements, and ultimately foster mutual comprehension and agreement on the potential issues threatening the validity and reliability of language tests. Contrary to the prevailing skepticism towards peer-to-peer feedback in higher education, the emergence of resistance and extended negotiation of meaning led pre-service teachers to understand the issues surrounding their test items more deeply and formulate suggestions to eradicate the problems effectively. Therefore, understanding and addressing the nature of resistance effectively is essential for educators at all levels. By identifying barriers to effective feedback, educators can adopt strategies to improve the quality of feedback exchanges. Addressing resistance not only enhances learning experiences by fostering a supportive and inclusive environment but also promotes student agency and ensures that all students feel valued and comfortable in showing engagement in the ongoing feedback process. Suffice it to say that this study highlights the necessity of creating authentic classroom environments where pre-service teachers can engage in reflective practices such as explanatory talk, negotiation, and constructive criticism to practice item construction and review, an essential phase in their future career as language teachers and assessors. The study also suggests implementing a flipped-classroom framework in ELTE courses, providing various learning opportunities for pre-service teachers to improve their assessment literacy and disciplinary content knowledge through feedback interactions, an often-overlooked aspect in language assessment and testing research and L2 teacher education.

**Implications**

Understanding the dynamics of preservice interactions and the emergence of resistance offers valuable implications for successfully implementing peer feedback in higher education contexts. Analyzing the sequential pattern of peer feedback interactions can help educators improve the effectiveness of feedback delivery and the reception process while enabling them to develop necessary strategies to address, mitigate, and manage resistance for a more collaborative learning atmosphere. Two benefits of peer feedback interaction are enhancing students' comprehension and improving test item quality. Additionally, the study highlights how crucial reflective teaching practices and peer feedback are to bettering language test design, item writing, and reviewing skills. This strategy prepares future educators for the demands of their profession, where tasks connected to assessment play a prominent role by encouraging dialogic reflections and peer review. According to this study, teacher education programs should be modified, and ELTE courses should be better adapted to suit the changing needs of educators in the 21st century.

**Recommendations**

This study initially recommends increasing the course hours devoted to foreign language testing, assessment, and evaluation courses for pre-service language teachers in the curriculum of foreign language education departments. Besides, designing the course in a flipped classroom format enhances pre-service teachers' chances to implement the theoretical concepts with real practice, prepare language exams, and write valid test

items. As practice is often a neglected aspect of this course due to time constraints and overcrowded classrooms, the implementation of peer feedback might serve as a solution to the aforementioned problems while also providing learners a chance to receive guidance, evaluation, and instruction on their language exams through item review and feedback interactions.

**Limitations**

The focus of the study has been the focal phenomenon of RTP in the four weeks of consecutive peer feedback sessions. However, the scope of the sessions included in the analysis chapter illustrates certain limits in the study. It should be noted that the presentations of the final versions of the language exam take place during the final week of summer school, session 8. Thus, more research on the phenomena of RTP and its efficiency against resistance practices might be explored deeply.

**Statements of Publication Ethics**

According to the authors, there are no unethical practices in this work that relate to publication ethics. The MA dissertation of the first author served as the basis for the investigation. Accordingly, the Ethics Commission of the university where the research was conducted during the 2019–2020 academic year granted ethical approval for the the study with the document number: 3322019 on October 04, 2019.

**Researchers' Contribution Rate**

Since the study was based on the first author's MA thesis, the first author was involved at every stage of the current study. The study also benefited greatly from the direction, guidance, and continuous feedback provided by the second author, who also served as the dissertation supervisor and dedicated mentor throughout this long journey.

| Authors | Literature review | Method | Data Collection | Data Analysis | Results | Conclusion |
|---|---|---|---|---|---|---|
| Author 1 | ☒ | ☒ | ☒ | ☒ | ☒ | ☒ |
| Author 2 | ☒ | ☐ | ☒ | ☐ | ☐ | ☐ |

**Conflict of Interest**

The writers disclose no conflict of interest.

**Acknowledgement**

The authors would like to express their sincere gratitude to the anonymous reviewers for their constructive feedback and valuable suggestions, which significantly contributed to the improvement of this manuscript. We would also like to extend our heartfelt thanks to the Scientific Research Projects Coordination Unit of Middle East Technical University for supporting the research that forms the basis of this study.

## REFERENCES

Badem-Korkmaz, F., Ekin, S., & Balaman, U. (2022). Pre-service language teachers' resistance to teacher trainer advice on task design for video-mediated L2 interaction. *Classroom Discourse*, 13(2), 212–230. https://doi.org/10.1080/19463014.2021.2020144

Bell, S. T., & Arthur Jr., W. (2008). Feedback acceptance in developmental assessment centers: The role of feedback message, participant personality, and affective response to the feedback session. *Journal of Organizational Behavior,* 29(5), 681–703. https://doi.org/10.1002/job.504

Beaumont, C., O'Doherty, M., & Shannon, L. (2011). Reconceptualizing assessment feedback: A key to improving student learning? *Studies in Higher Education,* 36(6), 671–687. https://doi.org/10.1080/03075071003731135

Bleakney, J., Mattison, M., & Ryan, J. (2019). "I was kind of angry": Understanding resistance to feedback in two tutor education courses. *Praxis: A Writing Center Journal*, *17*(1).

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22, 151–167. https://doi.org/10.1080/713695728

Boud, D., & Carless, D. (2018). Developing student feedback literacy: Helping students engage effectively with feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Boud, D., & Falchikov, N. (2007). *Rethinking Assessment in Higher Education: Learning for the Longer Term* (1st ed.). London: Routledge. https://doi.org/10.4324/9780203964309

Brown, H. D. (1996). Resources in language testing and assessment. *TESOL Quarterly, 30*(4), 781-781. https://doi.org/10.2307/3587936

Can Daşkın, N. (2017). *A conversation analytic study of reference to a past learning event in L2 classroom interaction: implications for informal formative assessment.* [Doctoral dissertation, Middle East Technical University].

Can, H. (2020). *A micro-analytic investigation into EFL teachers' test item reviewing interactions* [Unpublished doctoral dissertation, Middle East Technical University].

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Chiu, M. M. (2008). Flowing toward correct contributions during group problem solving: A statistical discourse analysis. *The Journal of the Learning Sciences*, *17*(3), 415-463. https://doi.org/10.1080/10508400802224830

Clayman, S. E. (2012). Turn-*Constructional Units and the Transition-Relevance Place*. In The Handbook of Conversation Analysis. https://doi.org/10.1002/9781118325001.ch8

Coleman, E. B. (1998). Using Explanatory Knowledge during Collaborative Problem Solving in Science. *The Journal of the Learning Sciences, 7*(3/4), 387–427. https://doi.org/10.1080/10508406.1998.9672059

Copland, F. (2010). Causes of tension in post-observation feedback in pre-service teacher training: An alternative view. *Teaching and Teacher Education*, *26*, 466–472. https://doi.org/10.1016/j.tate.2009.06.001

Copland, F., Ma, G., & Mann, S. (2009). Reflecting in and on post-observation feedback in initial teacher training on certificate courses. *English Language Teacher Education and Development, 12*, 14–23.

Cowan, J., & Creme, P. (2005). Peer assessment or peer engagement? Students as readers of their own work. *LATISS Learning and Teaching in the Social Sciences, 2*(2), 99-119. https://doi.org/10.1386/ltss.2.2.99/1

Crook, A., Mauchline, A., Maw, S., Lawson, C., Drinkwater, R., Lundqvist, K., Orsmond, P., Gomez, S., & Park, J. (2012). The use of video technology for providing feedback to students: Can it enhance the feedback experience for staff and students? *Computers & Education, 58*(1), 386–396. https://doi.org/10.1016/j.compedu.2011.08.025

Cushing, A., Abbott, S., Lothian, D., Hall, A., & Westwood, O. M. R. (2011). Peer feedback as an aid to learning – What do we want? Feedback. When do we want it? Now! *Medical Teacher*, *33*(2), e105-e112. https://doi.org/10.3109/0142159X.2011.542522

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, *83*(1), 1–27. https://doi.org/10.3102/0034654312474350

Falchikov, N. (2001). *Learning Together: Peer Tutoring in Higher Education*. Psychology Press.

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, *70*, 287–322. https://doi.org/10.2307/1170785

Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal, 81*(3), 285–300. https://doi.org/10.2307/329302

Goffman, E. (1964). The neglected situation. *American Anthropologist, 6*(2), 133–136.

Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of preservice English language teachers. *ELT Research Journal, 4*(2), 111-128. Retrieved from https://dergipark.org.tr/en/pub/eltrj/issue/28780/308006

Hatipoğlu, Ç. (2017), History of language teacher training and English language testing and evaluation (ELTE) education in Turkey. In Y. Bayyurt, & N. C. Sifakis (Eds.), *English Language Policies and Practices in the Mediterranean Countries and Beyond* (pp. 227-258). Peter Lang.

Hatipoğlu, Ç. (2023). Flipped spiral foreign language assessment literacy model (FLISLALM) for developing pre-service English language teachers' language assessment literacy. In D. Köksal, N. K. Ulutaş, & S. Arslan (Eds.), *Handbook of research on perspectives in foreign language assessment* (pp. 104-129). IGI Global. https://doi.org/10.4018/978-1-6684-5660-6

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Heaton, J.B. (1990*). Writing English language tests*. Longman.

Heritage, J. (1984a). *Garfinkel and ethnomethodology*. Polity Press.

Heritage, J., & Sefi, S. (1992). Dilemmas of Advice: Aspects of the Delivery and Reception of Advice in Interactions between Health Visitors and First Time Mothers. *Talk at Work,* 359-419.

Hepburn, A., & Potter, J. (2011). Designing the recipient: Managing advice resistance in institutional settings. *Social Psychology Quarterly, 74*(2), 216–241. https://doi.org/10.1177/0190272511408055

Hill, J., Berlin, K., Choate, J., Cravens-Brown, L., McKendrick-Calder, L., & Smith, S. (2021). Exploring the emotional responses of undergraduate students to assessment feedback: implications for instructors. *Teaching and Learning Inquiry: The ISSOTL Journal*, *9*(1), 294–316. https://doi.org/.20343/teachlearninqu.9.1.20

Hoey, E., & Kendrick, K. H. (2018). Conversation analysis. In A. M. B. De Groot & P. Hagoort (Eds.), *Research methods in psycholinguistics and the neurobiology of language: A practical guide* (pp. 151–173). Wiley. http://dx.doi.org/10.1002/9781394259762.ch8

Humă, B., Joyce, J. B., & Raymond, G. (2023). What Does "Resistance" Actually Look Like? The Respecification of Resistance as an Interactional Accomplishment. *Journal of Language and Social Psychology*, *42*(5-6), 497–522. https://doi.org/10.1177/0261927X231185525

Hyland, K. (2000). *Disciplinary Discourses: Social Interaction in Academic Writing.* London: Longman.

Kinnell, A. M. K., & Maynard, D. W. (1995). The delivery and receipt of safer sex advice in pretest counseling sessions for HIV and AIDS. *Journal of Contemporary Ethnography*, *24*(4), https://doi.org/10.1177/089124196024004002

Leyland, C. (2018). Resistance as a resource for achieving consensus: Adjusting advice following competency-based resistance in L2 writing tutorials at a British University. *Classroom Discourse*, *9*(3), 267–287. https://doi.org/10.1080/19463014.2018.1480966

Leyland, C. (2021). The interactional construction of the academic reader in writing tutorials for international students: An advice-giving resource. *Linguistics and Education, 61*, 100900. https://doi.org/10.1016/j.linged.2020.100900

Leung, S. (2002). Conflict talk: A discourse analytical perspective. *Studies in Applied Linguistics and TESOL*, *2*(3). https://doi.org/10.7916/salt.v2i3.1643

Liu, N.-F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education, 11*(3), 279–290. https://doi.org/10.1080/13562510600680582

Looney, A., Cumming, J., van der Kleij, F., van der Kleij, F., Harris, K. R., & Harris, K. R. (2017). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy & Practice, 25*(4), 1–26. https://doi.org/10.1080/0969594X.2016.1268090

McConlogue, T. (2020). *Assessment and Feedback in Higher Education: A Guide for Teachers.* UCL Press. https://doi.org/10.2307/j.ctv13xprqb

Metin, M. (2010). Biçimlendirici Değerlendirmeye Yönelik Öğretmen Adaylarinin Düşünceleri. *Milli Eğitim Dergisi, 40*(187), 293-310.

Muntigl, P. (2013). Resistance in couples counselling: Sequences of talk that disrupt progressivity and promote disaffiliation. *Journal of Pragmatics*, *49*(1), 18–37. https://doi.org/10.1016/j.pragma.2013.01.003

Ngar-Fun Liu & David Carless (2006) Peer feedback: the learning element of peer assessment, *Teaching in Higher Education, 11*:3, 279-290, https://doi.org/10.1080/13562510600680582

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education, 39*(1), 102–122. https://doi.org/10.1080/02602938.2013.795518

Nilson, L. B. (2003). Improving student peer feedback. *College Teaching*, *51*(1), 34–38. https://doi.org/10.1080/87567550309596408

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student-derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, *27*(4), 309–323. https://doi.org/10.1080/0260293022000001337

Park, I. (2014). Stepwise advice negotiation in writing center peer tutoring. *Language and Education*, *28*(4), 362–382. https://doi.org/10.1080/09500782.2013.873805.

Park, I. (2017). Questioning as advice resistance: Writing tutorial interactions with L2 writers. *Classroom Discourse.* https://doi.org/10.1080/19463014.2017.1307125

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128–138. https://doi.org/10.1016/j.tate.2019.05.003

Pekrun, R., Cusack, A., Murayama, K., & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction*, *29*, 115–124. https://doi.org/10.1016/j.learninstruc.2013.09.002

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4

Peräkylä, A. (2004). Conversation analysis. in C Seale, D Silverman, J Gubrium & G Gobo (eds)*, Qualitative Research Practice*. Sage, London, pp. 165-179.

Peräkylä, A. (2010). Shifting the perspective after the patient's response to an interpretation. *The International Journal of Psychoanalysis, 91*(6), 1363–1384. https://doi.org/10.1111/j.1745-8315.2010.00323.x

Pitt, E., & Norton, L. (2017). 'Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it. *Assessment & Evaluation in Higher Education*, *42*(4), 499–516. https://doi.org/10.1080/02602938.2016.1142500

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes. In J. M. Atkinson, & J. Heritage (Ed.), S*tructures of social action: Studies in conversation analysis*. 57–101. Cambridge University Press.

Pudlinski, C. (2002). Accepting and rejecting advice as competent peers: Caller dilemmas on a warm line. *Discourse Studies, 4*(4), 481–500. https://doi.org/10.1177/1461445602004004050l.

Pudlinski, C. (2012). The pursuit of advice on US peer telephone helplines: Sequential and functional aspects. In H. Limberg & M. A. Locher (Eds.), *Advice in Discourse* (pp. 232–252). John Benjamins.

Richardson, V. (2005). The diverse learning needs of students. In D. M. Billings & J. A. Halstead (Eds.), *Teaching in nursing* (2nd ed). Elsevier.

Robinson, S. J., Pope, D., & Holyoak, L. (2011). Can we meet their expectations? Experiences and perceptions of feedback in first-year undergraduate students. *Assessment & Evaluation in Higher Education*, *38*(3), 1–13. https://doi.org/10.1080/02602938.2011.629291

Ryan, T., & Henderson, M. (2018). Feeling feedback: Students' emotional responses to educator feedback. *Assessment & Evaluation in Higher Education, 43*(6), 880–892. https://doi.org/10.1080/02602938.2017.1416456

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144. https://doi.org/10.1007/BF00117714

Seckendorf, M. H. (1986). *Writing center conferences: An analysis* (Publication No. 8625994) [Doctoral dissertation, State University of New York at Albany].

Schegloff, E. A. (2007). *Sequence organization in interaction*. Cambridge University Press.

Sidnell, J., & Stivers, T. (2013). *The handbook of conversation analysis*. Blackwell

ten Have, P. (2001). Applied conversation analysis. In A. W. McHoul & M. Rapley (Eds.), *How to analyse talk in institutional contexts* (pp. 3–11). Continuum.

Topping, K., & Ehly, S. (Eds.). (1998). *Peer-assisted learning* (1st ed.). Routledge. https://doi.org/10.4324/9781410603678

Topping, K. J. (2009). Peer Assessment. *Theory Into Practice, 48*(1), 20–27. Published by Taylor & Francis, Ltd. Stable URL: https://www.jstor.org/stable/40071572

Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction, 20*(4), 339–343. https://doi.org/10.1016/j.learninstruc.2009.08.003

Tseng, S. C., & Tsai, C. C. (2010). Taiwan college students' self-efficacy and motivation of learning in online peer assessment environments. *Internet and Higher Education*, *13*(3), 164–169. https://doi.org/10.1016/j.iheduc.2010.06.002

Vehvilainen, S. (2001). Evaluative advice in educational counseling: The use of disagreement in the "stepwise entry" to advice. Research on Language & Social Interaction, *34*(3), 371–398. https://doi.org/10.1207/S15327973RLSI34-3_4

Vehviläinen, S. (2009) Problems in the Research Problem: Critical Feedback and Resistance in Academic Supervision, *Scandinavian Journal of Educational Research, 53*(2), 185-201, https://doi.org/10.1080/00313830902757592

Vehviläinen, S. (2012). Question-prefaced advice in feedback sequences of Finnish academic supervisions. In H. Limberg & M. A. Locher (Eds.), *Advice in Discourse* (pp. 31-52). Benjamins. https://doi.org/10.1075/pbns.221.04veh

Verkuyten, M. (2000). School Marks and Teachers' Accountability to Colleagues. *Discourse Studies*, *2*(4), 452–472. https://doi.org/10.1177/1461445600002004003

Waring, H. Z. (2005). Peer tutoring in a graduate writing centre: Identity, expertise, and advice resisting. *Applied Linguistics, 26*(2), 141–168. https://doi.org/10.1093/applin/amh041.

Waring, H. Z. (2007a). The multi-functionality of accounts in advice giving. *Journal of Sociolinguistics, 11*(3), 367–391. https://doi.org/10.1111/j.14679841.2007.00328

Waring, H. Z. (2007b). Complex advice acceptance as a resource for managing asymmetries. *Text & Talk – An Interdisciplinary Journal of Language, Discourse Communication Studies, 27*(1), 107-137. https://doi.org/10.1515/text.2007.005

Waring, H. Z. (2017). Going general as a resource for doing advising in post-observation conferences in teacher training. *Journal of Pragmatics, 110*, 20–33. https://doi.org/10.1016/j.pragma.2017.01.009

West, M. (2021). 'I'm Not Going to Tell You Cos You Need to Think About This': a Conversation Analysis Study of Managing Advice Resistance and Supporting Autonomy in Undergraduate Supervision. *Postdigital Science and Education*, *3*(1), 198–222. https://doi.org/10.1007/s42438-020-00194-5

Yang, Y.-F. (2016). Transforming and constructing academic knowledge through online peer feedback in summary writing. *Computer Assisted Language Learning, 29*(4), 683–702. https://doi.org/10.1080/09588221.2015.1016440

Yöney, R. Z. (2021). *Assuming learner behavior as an interactional resource in L2 testing and evaluation: Classroom interaction in a teacher education context*. [Master thesis, Hacettepe University]. https://doi.org/10.13140/RG.2.2.20584.34569