Applied Psycholinguistics (2026), **47**, e54, 1–19
doi:10.1017/S0142716426100976

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Features predicting item difficulty in English and Spanish elision task

Mónica Zegers[1,2], Julian M. Siebert[2], Lillian K. Durán[3], Javier Jasso[4], Amy S. Pratt[5], Jeremy Lichtmacher[6], Heather Murphy[2], Francesca Pei[2] and Maria Luisa Gorno-Tempini[2]

[1]Facultad de Ciencias Sociales, Universidad de Los Andes, Santiago, Chile, [2]Neurology, University of California San Francisco, USA, [3]College of Education, University of Oregon, USA, [4]Department of Spanish & Portuguese, The Ohio State University, USA, [5]Department of Communication Sciences and Disorders, University of Cincinnati, USA and [6]College of Medicine, Northeast Ohio Medical University, Rootstown, OH, USA

Corresponding author: Mónica Zegers; Email: mzegers@berkeley.edu

(Received 28 May 2025 UTC; revised 10 July 2026 UTC; accepted 27 July 2026 UTC)

Abstract

Phonological awareness (PA) is a foundational skill in early literacy acquisition across languages, including English and Spanish. While previous research has explored how factors such as word structure, phoneme characteristics, syllable stress, and syllable complexity influence PA item difficulty, the relative impact of these features—and how they may operate similarly or differently across languages—remains insufficiently understood. This study investigated the linguistic and task-format features that contribute to item difficulty in an elision task and compares these effects in English and Spanish. Using adaptive LASSO regression and dominance analysis, the study identified three significant predictors of item difficulty in English and six in Spanish. In both languages, the medial position of the deleted segment and the use of expressive response formats are the strongest predictors of increased item difficulty. In Spanish, items that produce nonwords involve longer phonological structures or require syllable deletion are associated with higher difficulty, whereas initial segment deletion tends to be easier. For English, word-level elision is associated with lower difficulty. Together, these findings highlight important considerations for designing more linguistically sensitive PA assessments in both English and Spanish.

Keywords: Assessment; elision; item difficulty; Spanish and English phonological awareness

Phonological awareness (PA)—the ability to distinguish and manipulate word sounds—is a critical cognitive skill that links early language development and reading acquisition (Lonigan, 2006; Melby-Lervåg et al., 2012). It consistently predicts word-reading outcomes in both English (Branum-Martin et al., 2012; Vellutino et al., 2004; Wolf et al., 2002; Ziegler & Goswami, 2005) and Spanish (Martínez & Goikoetxea, 2020; Míguez-Álvarez et al., 2022). Among the various

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

types of PA tasks (e.g., alliteration, blending, segmentation, and manipulation), elision tasks—which require deletion of specific sound units—show the strongest associations with reading outcomes (Anthony et al., 2011; Kilpatrick, 2012). In English-speaking children, elision tasks are predictive of reading and spelling performance through the early grades (Hogan, 2010; Hogan et al., 2005; Kilpatrick, 2012; Swank & Catts, 1994), and parallel findings in Spanish-speaking populations affirm their predictive value up to second grade (Anthony et al., 2011).

PA represents a unitary construct that manifests differently throughout development in both Spanish and English (Anthony & Francis, 2005; Anthony & Lonigan, 2004; Anthony et al., 2011; Schatschneider et al., 1999; Stanovich et al., 1984; Wagner & Torgesen, 1987), with stability increasing as children mature (Lonigan et al., 1998). Referred to as heterotypic continuity (see Anthony et al., 2011 for a review), it suggests that performance in PA assessments that emerge at different developmental stages is a manifestation of the same latent ability. PA develops from sensitivity to word-level units to manipulation of individual phonemes. Accordingly, PA assessments should be designed to capture this developmental progression by sampling skills across multiple levels of linguistic complexity.

Key structural differences exist between English and Spanish phonology and orthography, which may lead to distinct patterns in PA development and processing (Gottardo & Mueller, 2009). These languages differ in their regularity and degree of transparency of their writing systems for representing oral language. Spanish orthography is transparent and has fine phoneme–grapheme granularity, making PA easier to master (Goswami, 2010; Seymour et al., 2003; Wydell & Butterworth, 1999). On the other hand, English has a deep orthography, since graphemes can represent different phonemes, and phonemes may be represented by different graphemes. This is evident, for example, in the fact that the five vowel phonemes of Spanish are consistently represented by five vowel graphemes. English uses those same five graphemes to represent a considerably larger vowel inventory. Differences also show up in the available syllable structures and their relative distributions across both languages. While English phonology allows for complex syllable structures, Spanish phonology uses a comparatively simpler syllabic structure (i.e., preferring consonant-vowel syllables), which promotes syllable-level awareness and facilitates syllabic segmentation (Alvarez et al., 2001). These language-specific features may influence not only how children develop PA over time but also how they perform on elision tasks.

In addition to the evolving nature of the PA skill and the impact of specific language on its development, an additional challenge in measuring PA skills is that item features, such as phonological and lexical properties, influence item difficulty (Hogan, 2010; Scarborough, 2005). Prior research has seldom examined which specific word-level features together contribute to item difficulty, including receptive and expressive formats, in both English and Spanish. Existing studies largely focused on isolated features (e.g., only word length, or only phonological neighborhood), limiting generalizability and potentially omitting predictive variables and introducing bias. This study addressed the current gap in the literature by examining a comprehensive set of phonological and lexical features that influence item difficulty in elision tasks, with the goal of providing evidence that could advance the development of more sensitive elision tasks in both English and Spanish

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

by making item development more precise and efficient, targeting specific difficulty levels (Branum-Martin et al., 2012).

This study also adds to the literature on the assessment of elision in Spanish by focusing on Spanish-English bilinguals in the U.S. schooling context of primarily English instruction. The current body of work examining Spanish elision tasks has been conducted primarily in Spanish-speaking countries (e.g., Carrillo, 1994; Signorini, 1997), with a limited number in U.S. education contexts (Anthony et al., 2011; Branum-Martin et al., 2006). These linguistic contexts differ significantly, given differences in language of instruction, and quantity and quality of exposure to both languages across natural settings.

Item features

Response modality: Receptive or expressive

Receptive elision tasks present response options as pictures and require children to identify the image that corresponds to the target word (e.g., "What is 'snowman' without 'snow'?" with picture choices of a man, a pan, and a cat). In contrast, expressive elision tasks require children to generate the correct response verbally, without the support of visual cues.

Past studies examined differences between receptive and expressive modalities and found differences in task difficulty. One study of a Spanish blending task found that expressive items are more challenging than a parallel receptive version (Anthony et al., 2011). The difficulty of expressive tasks may stem from their reliance on the coordination of phonological retrieval, working memory, vocabulary knowledge, and speech production (Baddeley, 2003), which can impose greater cognitive and linguistic demands.

Both receptive and expressive tasks have their own merits. Schatschneider et al. (1999) reported that English expressive blending and elision tasks are most effective in distinguishing children's PA abilities, whereas certain receptive tasks, such as first sound comparison, are easier and better suited for children with lower PA skills. Receptive tasks are often used to assess the precision of phonological representations, particularly in children with speech sound disorders, for whom speech motor difficulties may mask representational skills (Anthony et al., 2010), and for younger children with emerging PA skills (Anthony et al., 2011; Farver et al., 2007; Schatschneider et al., 1999).

Elision level: Word, syllable, phoneme

Psycholinguistic grain size theory, where grain sizes refer to the units of varying lengths (e.g., phonemes, onsets, and syllables) that are manipulated in real time (Goswami, 2010), provides a useful framework when exploring item difficulty. Children exposed to different languages make use of different phonological units reflecting specific phonological structures and orthographic mappings of the particular language. When children first start learning one or more languages, they process words and routine phrases as whole, unanalyzed units in a broad, holistic way; instead of focusing on individual sounds, they rely on suprasegmental acoustic

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

cues, such as changes in pitch and duration, to recognize words and syllables (Anthony & Francis, 2005).

As their vocabulary grows, they gradually fine-tune these representations, becoming more sensitive to the subtle differences between sounds (Anthony et al., 2010). This progression is reflected on how children perform on elision tasks, which may require removing linguistic material at various levels: whole words, syllables, onset-rime, and phonemes. The smaller the unit, the harder it is to manipulate, requiring more precise PA (Anthony & Francis, 2005). Lonigan et al. (1998) found that while many young children could manipulate whole syllables, fewer could manipulate smaller units, such as phonemes, suggesting that syllable awareness lays the groundwork and is a prerequisite for more advanced phonological skills.

This developmental progression from syllable to phoneme has also been observed in Spanish-speaking children. Accuracy and difficulty in monolingual Spanish-speaking children go from manipulation of syllables (e.g., gi-ra-sol) to onset-rime (s-ol), and finally to discrete phonemes (s-o-l; Goikoetxea, 2005). Preliterate Spanish-speaking children manipulate syllables more easily than phonemes (Jiménez, 1992), and syllable manipulation has been found to be easier than phoneme manipulation (Jiménez González & Ortiz Gonzalez, 1994). These findings underscore the influence of children's gradual refinement of phonological representations on task difficulty, reinforcing the notion that PA develops progressively, with children mastering larger sound units before smaller ones, which is congruent with the psycholinguistic grain size theory (Lallier & Carreiras, 2018).

Elision position: Initial, medial, final

Item difficulty is also influenced by the position of the deleted sound (i.e., initial, medial, or final) in elision tasks (Chafouleas et al., 2001). In English, medial positions are more difficult to identify or manipulate than initial or final positions (McBride-Chang, 1995), a finding attributed to differences in memory demands and articulation. The cognitive interpretation highlights primacy and recency effects in short-term memory, which favor stimuli in initial and final positions, while disfavoring those in the middle. In the articulatory explanation, possible coarticulation interference may impact medial speech sounds to a greater degree, given that both the preceding and following phonemes affect the perception of medial phonemes (Stage & Wagner, 1992).

Spanish-speaking children also show positional effects, although the literature conflicts. Carrillo (1994) and Fumagalli et al. (2010) reported that it is easier to accurately respond to PA items targeting the initial segments than those targeting the medial or final segments. Goikoetxea (2005), on the other hand, reported that, while initial units are easier for literate children, ending units are easier for preliterate children. Additional research is needed to clarify these discrepancies and to expand to multilingual children.

Finally, in addition to purely positional considerations, differences in task performance across languages could be expected due to the phonotactics of each language. One clear difference is in the relative absence/presence of coda consonants. English phonology is largely permissive of syllable coda consonants, and various sound classes can and routinely do appear in coda position (stops like /p t/

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

and fricatives like /f/ /z/ in words like "tub" and "huff"). In contrast, Spanish shows more restrictions on which consonants can appear in the coda position. Specifically, consonants that appear in the coda are primarily /s/, /n/, /r/, /l/, and /d/ (e.g., "rosas," "sal," and "verdad"). In fact, there appear to be both coda-weakening and syllable restructuring processes to keep non-coda (CV) syllables when possible (e.g., las hadas: /las.ˈa.das/ → [la.ˈsa.ðas]; verdad: /beɾ.ˈdad/ → [beɾ.ˈðað]; Hualde, 2005; Hualde et al., 2020). In sum, differences in the phonotactics of both languages may produce differences in item difficulty. Overall, these findings suggest that, while elision task difficulty is affected by position in both English and Spanish, language-specific phonological features modulate how these effects manifest across developmental stages.

Target lexicality: Word or nonword

Phonological representations in memory help individuals store and access speech sound information, including a word's meaning, spelling, and articulation (Anthony et al., 2010). Young children (e.g., kindergarten, first grade) exhibit a lexicality effect, as they generally perform better on PA tasks that use real words rather than nonwords. This is likely because they are familiar with words they have heard before, which are entrenched in their memory (Metsala, 1999). At the same time, there is evidence that the same underlying phonological processing abilities are recruited regardless of lexicality. Children's syllable blending and elision using real words (such as compound words: sunflower take away sun is flower) are associated with the same operations while using nonword syllables and phonemes, even after accounting for their oral language skills (Lonigan et al., 1998). So, although real high-frequency words may confer an advantage, both real and nonwords appear to use the same processes.

Phonological neighborhood and word frequency

Lexical features such as phonological neighborhood density and word frequency influence PA and phonological processing. *Phonological neighborhood density* refers to the number of words that differ from a target word by a single phoneme (e.g., *pin*: bin, win, pen, pit; Charles-Luce & Luce, 1990; Dollaghan, 1994). According to lexical restructuring theory, as children's vocabularies grow, their phonological representations become more detailed and segmental, becoming more proficient in distinguishing between multiple similar-sounding words for words in dense neighborhoods (Charles-Luce & Luce, 1990).

The evidence for the role of neighborhood density in PA tasks is mixed. Some studies find that dense neighborhoods facilitate performance by providing a more fine-tuned phonological contrast (De Cara & Goswami, 2003; Metsala, 1999). Others report increased errors or slower reaction times in dense onset neighborhoods, possibly due to increased lexical competition during processing (Foy & Mann, 2009). These conflicting findings may reflect interactions with word frequency, a factor known to influence lexical access and phonological representation. For example, Hogan et al. (2011) found that high word frequency offsets the effects of neighborhood density, suggesting that robust lexical representations may reduce reliance on neighborhood cues.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Word frequency itself has been consistently linked to PA performance. High-frequency words are generally processed more accurately, likely because repeated exposure leads to stronger, more stable phonological representations (Metsala, 1999; Swan & Goswami, 1997). In contrast, low-frequency or phonologically complex words tend to have underspecified representations, making them harder to process. Research suggests that sound blending is easier with high-frequency than with low-frequency words (Roth et al., 2006; Troia et al., 1996).

Integrating these findings, the lexical restructuring model (Metsala & Walley, 1998) proposes that both phonological neighborhood density and word frequency shape the development of PA, potentially through interactive mechanisms. Hogan et al. (2011), for example, found that words with denser neighborhoods lead to more accurate responses only with low-frequency words. These results support a threshold hypothesis: either high word frequency or high neighborhood density alone may be sufficient to support accurate phoneme awareness. Together, these findings underscore the importance of considering lexical features when designing PA tasks. This is particularly relevant for bilingual or language-minority children, whose lexical access (and thus the influence of frequency and neighborhood density) may vary across languages (Goodrich & Lonigan, 2016).

Word length

Another factor that influences children's PA performance across modalities and languages is word length. In English (Treiman & Weatherston, 1992), preschoolers and kindergarteners have more difficulty isolating the initial consonant of longer words compared to shorter words. Similarly, older elementary children respond less accurately as nonword length increased in a phoneme elision task (Sasisekaran & Byrd, 2013). In Spanish, children perform better on a PA isolation task when the first consonant belongs to a shorter rather than a longer word (Jiménez González & García, 1995). This pattern is also observed in children completing syllable and phoneme oddity tasks in Greek, a transparent language like Spanish, who show better performance with short than with long words (Aidinis & Nunes, 2001). A possible explanation for this result is the increased taxation on phonological short-term memory for longer words (Perrachione et al., 2017). However, while the effect of word length is consistently observed across languages, the unit of measurement varies. Jiménez González and García (1995) used the number of syllables as a criterion for length because Spanish has a well-defined syllabic structure that makes it easier to distinguish between monosyllabic or shorter words and trisyllabic or longer words. For English, it has been suggested that the number of phonemes is a more informative metric that explains children's performance (Treiman & Weatherston, 1992), as children perform better with two-phoneme one-syllable words than with three-phoneme one-syllable words.

Study aim

This study investigates a comprehensive set of phonological, lexical, and task-format features that influence item difficulty in parallel English and Spanish elision tasks. Understanding which factors make items easier or harder is of great benefit to

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

authors of such assessments because the intentional manipulation of these features allows for a targeted and efficient process of developing easy or hard items for a specific purpose.

We leverage our robust item-response theory calibrations of these tasks, which drew on large samples, representative of the California public-school student population. In addition, using a multivariate analysis approach allows us to address the potential omitted-variable bias present in prior research, which often focused on isolated features, and helps advance our understanding of the complex relationships between item features and difficulty.

Method

Measure development and calibration

The data for this study were drawn from a larger multi-year project that used purposive sampling to reflect the demographic characteristics of California's public-school population, including race/ethnicity and English language acquisition status. English measures were calibrated using a sample of 1,950 kindergarten through second-grade monolingual and multilingual students attending schools where instruction was delivered in English. Spanish measures were calibrated using a sample of 1,810 Spanish–English bilinguals in kindergarten through second grade, enrolled in either English-only or dual-language programs. Additional details on measure development, calibration procedures (including calibration samples), and validity evidence are provided in Appendix A.

Item features

Below we list various item features than can impact the Rasch item location, also known as the difficulty parameter. Item descriptive statistics are provided on Table 1.

Difficulty

Item difficulty refers to a test item's location (b) estimate, as obtained from a Rasch model fit to the response data using maximum likelihood estimation (Rasch, 1980). Rasch modeling allows for both item difficulty and student ability to be represented on the same (logit) scale for the latent trait being measured. Then, for dichotomously scored items, an item's difficulty/location corresponds to the ability level at which a student has an equal probability of responding correctly or incorrectly to said item. As a result, more demanding items have higher item location estimates.

Response modality

Elision can be assessed in a receptive or expressive format; to avoid confusion or response-mode switching, we administered the two types of tasks separately. In the receptive task, children saw three images and heard an audio labeling the images (e.g., *"fox," "box," and "onion"*) followed by the prompt (e.g., "*The word is 'mailbox.' What is 'mailbox' take away 'mail'?")*. Children were required to select the image

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Table 1. Item descriptive statistics

	Spanish		English	
Variable	Count	%	Count	%
Measure type	163		157	
Expressive	114	69.94	108	68.79
Receptive	49	30.06	49	31.21
Elision level	163		157	
Word	26	15.95	41	26.11
Syllable	71	43.56	54	34.39
Phoneme	69	42.33	65	41.40
Elision position	163		157	
Initial	74	45.40	60	38.22
Medial	57	34.97	49	31.21
Final	32	19.63	48	30.57
Target lexicality	163		157	
Nonword	24	14.72	28	17.83
Word	139	85.82	129	82.17

Variable	*n*	*M*	SD	Min	Max	*n*	*M*	SD	Min	Max
Difficulty	163	1.36	1.80	−1.67	6.27	157	−0.14	2.30	−4.45	5.52
Initial word										
PTAF	93	56.13	152.36	0.99	1289.27	110	83.00	161.45	0.33	954.82
PTAN	118	4.03	5.59	0	32	148	8.22	12.78	0	56
Freq. per mill.	163	52.38	220.09	0.00	2331.37	152	42.25	108.61	0.06	796.65
Phono. length	163	6.89	2.02	4	12	157	5.43	1.77	2	10
Syllable number	163	3.02	0.89	2	5	157	1.94	0.82	1	5
Target result										
PTAF	127	171.75	337.81	1.96	2565.76	126	305.28	442.22	1.07	2294.61
PTAN	128	12.08	7.68	0	36	127	30.53	16.80	0	62
Freq. per mill.	139	207.24	931.49	0.02	8952.76	127	511.92	1352.05	0.45	9773.41
Phono. length	163	4.47	1.02	3	7	157	3.40	1.30	1	7

Note: We deliberately avoided retrieving information on neighborhood and frequency from multiple databases, as doing so could introduce inconsistencies due to differences in estimation methods.

corresponding to the target word. Each trial showed the correct response (e.g., "box"), one phonological foil (e.g., "fox"), and one unrelated foil (e.g., "onion"). Items for the receptive task prioritized imageability and concreteness, with accompanying pictures designed to be universally recognizable and culturally unbiased.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

For the expressive task, children heard a prompt (e.g., "*The word is sunglasses. What is 'sunglasses' take away 'sun'?*") and verbally responded with the remaining segment of the word. The initial word for the expressive task was always a real word, but the target could be either a real word or a nonword.

Elision level: Word, syllable, phoneme

Items were designed to target elision at the word (in compound words), syllable, and phoneme levels. For six items (three in English and three in Spanish), an elided vowel resulted in both syllable and phoneme elision (e.g., "*What is 'along' takeaway /ə/?*"). Consequently, these examples were categorized as both phoneme and syllable elision.

Elision position: Initial, medial, final

Items were classified as initial, medial, or final. Initial came from the beginning of the word (e.g., "*What is 'leg' takeaway /l/?*"), medial from the middle of the word (e.g., "*What is 'yesterday' takeaway 'ter'?*"), and final from the end of the word (e.g., "*What is 'sheep' takeaway /p/?*"). For word-level elision, the elision position was either initial or final.

Initial word features

Information about the initial words' phonological length, neighborhood density, and frequency per million was obtained from the CLEARPOND database (Marian et al., 2012). Phonological length refers to the number of phonemes that make up a word. Phonological neighborhood size(PTAN) refers to the number of phonological neighbors through substitution, addition, and deletion. Phonological neighborhood frequency (PTAF) is the average frequency of a word's orthographic and phonological neighborhood. Finally, frequency per million corresponds to the word's lexical frequency per million words in the CLEARPOND corpus. Not all items had information in CLEARPOND, especially Spanish words and longer (and likely, less frequent) items.

Target features

Targets (i.e., the resulting string after completion of the elision operation) were classified as words or nonwords. For example, "*What is table take away /t/?*" results in the real word "able"; "*What is watermelon take away 'me'*" results in the nonword "waterlon". Information about length, neighborhood, and frequency of the targets was also collected using the CLEARPOND database. Item descriptive statistics are available in Table 2.

Analytical strategy

We implemented a three-step analytic procedure to identify and evaluate the predictors of item difficulty. First, we employed an adaptive Least Absolute Shrinkage and Selection Operator (LASSO; Tibshirani, 1996; Zou, 2006) regression with 10-fold cross-validation for penalty parameter (λ) tuning on pre-standardized variables. This method uses weights based on the data to apply different levels of penalty to each predictor, reducing the influence of weak or unimportant variables and pushing their

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Table 2. Dominance analysis results for LASSO-selected variables for English

	Dominance analysis		
LASSO-selected item characteristics	Rank	% of R2	OLS regression coef. (B)
Position: Medial	1	0.4901	2.839 (0.178)***
Measure: Expressive	2	0.3181	1.846 (0.146)***
Elision: Word	3	0.1918	−1.125 (0.144)***
MSE	0.646		
***R*-sq.**	0.877		
Out-of-sample *R*-sq.	0.872		
Obs.	157		

Note: Robust standard errors in parenthesis; ***p <0.01, **p <0.05, *p <0.1. Position-Medial: The deleted syllable or phoneme occurs in the middle of the word. Measure-Expressive: The item required participants to verbally produce the remaining segment of the word or pseudoword. Elision-Word: The deleted segment was an entire word within a compound word.

effects to zero, which helps remove variables that add little meaningful information. Second, we entered the selected predictors into an ordinary least squares regression to obtain unpenalized coefficient estimates and corresponding standard errors, clarifying the direction and magnitude of each predictor's association with item difficulty. Lastly, we conducted a dominance analysis (Azen & Budescu, 2003) to quantify and rank the relative importance of the retained predictors. The resultant standardized dominance weights sum to 1 and can be interpreted as the percentage each variable contributes to the total variance explained. This facilitates direct comparison of predictors' contributions to the overall model fit. Given that not all item features could be obtained or computed for all items (e.g., phonological neighborhood frequency or frequency per million for targets that are nonwords), we carried out the aforementioned analysis steps twice for each language: The first set draws on the complete set of items, foregoing predictors that could not be obtained for the full item set, and is presented in the main text. The second set of analyses draws on the subset of items for which we could obtain information on all item features, systematically excluding those items where the target is a nonword, and can be considered a sensitivity analysis. It is therefore included in Appendix B (Tables S2 and S3). All analyses were carried out using STATA/MP 17.0 (StataCorp, 2023).

Results

Adaptive LASSO regression

The main adaptive LASSO models included all available items (English: $N = 157$; Spanish: $N = 163$) but were limited in the number of predictors to those variables that could be computed for nonword targets.

English results

The English model (see Table 2) selected the medial position, the expressive method of response, and elision at the word level as relevant predictors of item difficulty.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Table 3. Dominance analysis results for LASSO-selected variables for Spanish

	Dominance analysis		
LASSO-selected item characteristics	Rank	% of R2	OLS regression coef. (B)
Position: Medial	1	0.3586	1.746 (0.179)***
Measure: Expressive	2	0.3058	1.579 (0.150)***
Position: Initial	3	0.1408	−0.279 (0.144)
Target lexicality: Word	4	0.1216	−0.578 (0.240)***
Target phono. Length	5	0.0419	0.125* (0.070)
Elision: Syllable	6	0.0312	0.533 (0.124)***
MSE	0.6		
***R*-sq.**	0.813		
Out-of-sample *R*-sq.	0.794		
Obs.	163		

Note: Robust standard errors in parenthesis; ***p <0.01, **p <0.05, *p <0.1. Position-Medial: The deleted syllable or phoneme is located in the middle of the word. Measure-Expressive: The item required participants to verbally produce the remaining segment of the word or pseudoword. Position-Initial: The deleted word, syllable, or phoneme is located at the beginning of the word. Target Lexicality-Word: Target item after elision results in a real word. Target Phono. Length: Phonological length of the target item after elision. Elision-Syllable: The deleted segment was a syllable within a word.

The same feature selection was made by the alternative model (Appendix B, Table S1).

Spanish results

Like English, the Spanish model (see Table 3) also selected the medial position of the deleted section and the expressive response method, but differed in that it included initial elision position and elision at the syllable level, while excluding elision at the word level. The Spanish model also incorporated target lexicality and phonological length. The same feature selection was made by the alternative model, which uses additional target-level predictors but excludes items with nonword targets (Appendix B, Table S2).

Dominance analysis

English results

For English, medial position was the most dominant predictor of item difficulty, accounting for approximately 49.0% of the overall variance explained by the model (see Table 2). Items involving medial deletion were associated with greater difficulty. The expressive response format ranked second in dominance, accounting for about 31.8% of the explained variance. Items requiring expressive responses were associated with increased item difficulty. Elision at the word level ranked third, accounting for roughly 19.2% of the explained variance. Word-level elision was associated with lower difficulty, suggesting that items requiring the deletion of an entire word within a compound were easier to respond to accurately than items involving other types of elision.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Overall, the English results indicate that item difficulty is primarily driven by the position of deletion and the response format, with elision level (i.e., word) playing a more limited role and going in the opposite direction.

Spanish results

In Spanish, medial position was again the most dominant predictor, accounting for approximately 35.9% of the overall variance explained (see Table 3). As in English, medial deletion was associated with greater difficulty. The expressive response format also ranked second, accounting for about 30.6% of the explained variance and being associated with higher difficulty.

Initial position for deletion ranked third in the dominance analysis, accounting for approximately 14.1% of the explained variance. Initial deletion was associated with lower difficulty, indicating that items requiring deletion at the beginning of the word were easier to respond to accurately than items involving medial or final deletion. Target lexicality ranked fourth, accounting for about 12.2% of the explained variance. Items that resulted in real words after deletion were associated with lower difficulty compared to items resulting in nonwords.

Phonological length of the target word was fifth in the dominance analysis, positively affecting item difficulty, and accounting for approximately 4.2% of the explained variance. Finally, syllable-level elision ranked sixth, accounting for about 3.1% of the explained variance, and was associated with greater difficulty.

The relatively small proportions of variance explained by phonological length and syllable-level elision indicate that these features play a limited role in accounting for item difficulty compared to positional and response format factors.

Discussion

Although elision tasks are commonly used to assess PA skills, there is little published information about which item features influence the item difficulty of these tasks. Our results address this gap and identify the variables explaining item difficulty in elision tasks used with monolingual and multilingual children living in the United States and attending kindergarten to second grade. These analyses are valuable to inform item development in each language. The findings contribute to the published research, which has largely focused on Spanish-speaking monolingual children and has not provided side-by-side comparisons of Spanish and English. Below, we first discuss the strongest predictors across both language tasks and the findings unique to each language.

Factors influencing difficulty in both English and Spanish

The two strongest predictors of item difficulty across languages were medial deletion and expressive response format. Together, they accounted for 80.8% of the variance explained by the model in English and 66.4% in Spanish.

The selection of the medial deletion variable aligns with previous research showing that phonemes in medial positions (e.g., "*What is 'belt' takeaway '/l/'?*") are more challenging to process than those at word edges (McBride-Chang, 1995; Stage

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

& Wagner, 1992). One explanation is that medial segments are subject to stronger coarticulation effects, as they are acoustically and articulatorily influenced by both preceding and following segments, which can obscure clear phonetic cues. In addition, medial segments (for both syllable and phoneme-level elision items) tend to be less perceptually salient than initial or final segments, which benefit from prosodic and boundary-related cues. Reduced salience, in turn, may place greater demands on phonological short-term memory, making it more difficult to accurately process, maintain, and retrieve medial phonological information (McBride-Chang, 1995; Stage & Wagner, 1992).

The second most influential variable affecting item difficulty across languages was response format: expressive tasks were consistently more difficult than receptive tasks. Receptive tasks place lower cognitive and linguistic demands on participants because responses are presented visually and selected from a constrained set of options (in this measure, one correct response, one phonological foil, and one unrelated foil), eliminating the need for independent language production. In contrast, expressive tasks require the retrieval, manipulation, and articulation of phonological information without external support, placing greater demands on phonological working memory and speech production processes. These demands may be especially pronounced for young children and for children with language or memory difficulties (Anthony et al., 2010, 2011; Farver et al., 2007; Schatschneider et al., 1999). Additionally, in English tasks, expressive responses may be particularly challenging for emerging bilinguals who are still developing proficiency and are less familiar with the lexical items used. In such cases, maintaining unfamiliar or weakly specified phonological representations in working memory while simultaneously extracting and manipulating the relevant units likely increases cognitive load. Although this result was expected, there is relatively little published work directly documenting the impact of response format on elision task difficulty.

Neither phonological neighborhood density nor word frequency emerged as key predictors of difficulty in either model. The relationship between these variables and task difficulty also shows conflicting evidence in previous research (De Cara & Goswami, 2003; Foy & Mann, 2009; Hogan et al., 2011; Metsala, 1999). It is possible that the strength of the correlation is moderated by developmental stage or measure design. An alternative explanation is that previous studies did not consider all the variables analyzed in this study simultaneously. This may have introduced omitted-variable bias, making it difficult to isolate the true influence of lexical and phonological properties on difficulty levels. The inclusion of multiple interacting variables in the present study may have better accounted for the complexity of PA tasks, potentially explaining discrepancies with earlier findings. More research using a similar set of variables is needed to support this claim.

Factors influencing difficulty in English

In addition to the trends across languages mentioned above, we also observed language-specific findings. In English, word-level segmentation was associated with lower difficulty, consistent with a developmental trajectory in which word awareness emerges before more complex phonemic manipulations. This aligns with research suggesting that children initially process words as whole units before

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

developing the ability to manipulate smaller phonological segments (Treiman & Zukowski, 1991). The findings support the notion that PA develops progressively, with word-level manipulations being relatively easier than finer phonemic ones by school age (Anthony & Francis, 2005).

Factors influencing difficulty in Spanish

For Spanish, the deletion of initial units was associated with lower difficulty. This result is in line with findings with Spanish monolingual children (Carrillo, 1994; Fumagalli et al., 2010; also, Goikoetxea, 2005, for literate children).

Items that resulted in real word targets were associated with lower difficulty, consistent with evidence that lexical familiarity supports phonological manipulation by providing more stable and well-specified memory representations (Metsala, 1999; Lonigan et al., 1998). In contrast, items yielding longer phonological targets were associated with greater difficulty, likely reflecting increased memory and sequencing demands involved in maintaining and manipulating longer sound strings. This finding aligns with existing research showing that, as length increases, performance accuracy in PA tasks decreases, both in Spanish (Jiménez González & García, 1995) and other alphabetic languages (Aidinis & Nunes, 2001). Although these effects accounted for a smaller proportion of variance than medial deletion and expressive format, together they suggest that both the quality of memory representations and memory load contribute to elision task difficulty. Another possible explanation is that word length may play a different role in Spanish and English because of differences in their phonological systems. English has a larger phoneme inventory, allowing many words to be distinguished with relatively short forms. Spanish has a smaller phoneme inventory and often relies on longer words to distinguish meaning. As a result, variation in word length may be more noticeable in Spanish, making word length a stronger contributor to task difficulty.

Finally, syllable-level elision also emerged as an important predictor of item difficulty. This result was not surprising, as prior research has found that tasks requiring syllable identification or manipulation are highly informative for capturing individual differences in PA in Spanish, often more so than phoneme-level elision tasks (Anthony et al., 2011). This suggests that syllable-level elision may benefit from a "Goldilocks effect" in its sensitivity to student variability when compared to word and phoneme-level elision. Word-level elision, which involves identifying free morphemes in compound words, may have been too easy, resulting in high accuracy and limited variance. In contrast, phoneme-level elision may have been too difficult, leading to low accuracy across participants. In both cases, restricted variability would reduce the items' utility in distinguishing phonological skill levels. Manipulation at the syllable level may strike an optimal balance.

An unexpected finding, however, was the directionality of this effect, with syllable-level elision associated with greater difficulty in Spanish. This finding was not congruent with Anthony and Francis's (2005) assertion that children exposed to languages with salient syllabic structures tend to develop syllable awareness more easily. Spanish has comparatively regular syllable structure and clear syllable boundaries, properties that have been linked to syllable-based speech segmentation (Cutler et al., 1986). Moreover, Cutler et al. (1986) found that children tend to apply

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

their dominant segmentation routine across languages, even when it is not optimally suited to the language being processed. In the present context, this pattern may reflect the combined influence of English or Spanish in the bilingual phonological development process found in the U.S. and children's instructional experiences in predominantly English-only contexts.

Bilingual children learning both English and Spanish may follow phonological trajectories that differ from those of Spanish-speaking monolinguals (Melby-Lervåg & Lervåg, 2011), with additional variation linked to language dominance. Instructional context may further shape these outcomes. Approximately 80% of the Spanish-speaking students in the calibration sample were enrolled in English-only instructional programs, where PA instruction often emphasizes phoneme segmentation, such as the use of rimes in early elementary (Gottardo & Mueller, 2009; i.e., isolating the initial sounds in l-og, d-og, etc.). Because English and Spanish are both alphabetic languages and share many overlapping letter-sound correspondences (e.g., "s" is /s/ in "sun" and "sal"; "ch" is /tʃ/ in "chin" and "chamarra"), students may transfer phonological strategies acquired in English to Spanish tasks. Such transfer could increase phonemic sensitivity while simultaneously reducing efficiency in syllable segmentation, particularly if instruction consistently emphasizes phonemes. Future work examining children's error patterns may help evaluate this explanation. Overall, the practical impact of this pattern appears limited, as syllable-level elision accounted for a relatively small proportion (3.1%) of the overall variance explained by the model.

Limitations and future research

While the study accounted for multiple phonological and lexical features simultaneously, the list of features is not exhaustive. Further research is needed to explore the influence of additional features, such as biphone probability (Goodrich & Lonigan, 2016) and syllable-initial consonant clusters (Jiménez González & García, 1995).

In addition, future research should explore item difficulty estimates in both English and Spanish for bilingual students, while accounting for their language background, language proficiency, and language of instruction to better understand how these factors shape phonological processing. This information will be critical in refining assessments and enhancing the early identification of reading difficulties that are less confounded by environmental factors.

Conclusion

This study offers valuable insights into the phonological, linguistic, and task-level features that influence item difficulty in English and Spanish elision tasks. The findings highlight medial deletion and expressive response format as dominant contributors increasing task difficulty, underscoring the importance of phonological processing demands in both languages. The results also reveal language-specific patterns, with elision at the word-level contributing to English item difficulty and elision at the syllable-level contributing to item difficulty in Spanish.

From an assessment development perspective, these results have important implications for refining elision tasks in English and Spanish. Test developers can

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

more accurately and systematically create items of varying difficulty levels, using empirical guidance on what makes an item more or less difficult.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0142716426100976

References

Aidinis, A., & Nunes, T. (2001). The role of different levels of phonological awareness in the development of reading and spelling in Greek. *Reading and Writing*, **14**, 145–177.

Alvarez, C. J., Carreiras, M., & Taft, M. (2001). Syllables and morphemes: Contrasting frequency effects in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **27**(2), 545.

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, **14**(5), 255–259.

Anthony, J. L., & Lonigan, C. J. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology*, **96**(1), 43.

Anthony, J. L., Williams, J. M., Aghara, R. G., Dunkelberger, M., Novak, B., & Mukherjee, A. D. (2010). Assessment of individual differences in phonological representation. *Reading and Writing*, **23**, 969–994.

Anthony, J. L., Williams, J. M., Durán, L. K., Gillam, S. L., Liang, L., Aghara, R., Swank, P. R., Assel, M. A., & Landry, S. H. (2011). Spanish phonological awareness: Dimensionality and sequence of development during the preschool and kindergarten years. *Journal of Educational Psychology*, **103**(4), 857.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, **8**(2), 129–148. https://doi.org/10.1037/1082-989X.8.2.129

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, **36**(3), 189–208.

Branum-Martin, L., Mehta, P. D., Fletcher, J. M., Carlson, C. D., Ortiz, A., Carlo, M., & Francis, D. J. (2006). Bilingual phonological awareness: Multilevel construct validation among Spanish-speaking kindergarteners in transitional bilingual education classrooms. *Journal of Educational Psychology*, **98**(1), 170.

Branum-Martin, L., Tao, S., Garnaat, S., Bunta, F., & Francis, D. J. (2012). Meta-analysis of bilingual phonological awareness: Language, age, and psycholinguistic grain size. *Journal of Educational Psychology*, **104**(4), 932.

Carrillo, M. (1994). Development of phonological awareness and reading acquisition: A study in Spanish language. *Reading and Writing*, **6**, 279–298.

Chafouleas, S. M., VanAuken, T. L., & Dunham, K. (2001). Not all phonemes are created equal: The effects of linguistic manipulations on phonological awareness tasks. *Journal of Psychoeducational Assessment*, **19**(3), 216–226.

Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, **17**(1), 205–215.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, **25**(4), 385–400.

De Cara, B., & Goswami, U. (2003). Phonological neighbourhood density: Effects in a rhyme awareness task in five-year-old children. *Journal of Child Language*, **30**(3), 695–710.

Dollaghan, C. A. (1994). Children's phonological neighbourhoods: Half empty or half full? *Journal of Child Language*, **21**(2), 257–271.

Farver, J. M., Nakamoto, J., & Lonigan, C. J. (2007). Assessing preschoolers' emergent literacy skills in English and Spanish with the Get Ready to Read! screening tool. *Annals of Dyslexia*, **57**, 161–178.

Foy, J. G., & Mann, V. A. (2009). Effects of onset density in preschool children: Implications for development of phonological awareness and phonological representation. *Applied Psycholinguistics*, **30**(2), 339–361.

Fumagalli, J., Wilson, M., & Jaichenco, V. (2010). Sensibilidad a la información fonológica en niños lectores y prelectores hispanoparlantes. *Revista Neuropsicología Latinoamericana*, **2**(1), 68–77.

Goikoetxea, E. (2005). Levels of phonological awareness in preliterate and literate Spanish-speaking children. *Reading and Writing*, **18**, 51–79.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Goodrich, J. M., & Lonigan, C. J. (2016). Lexical characteristics of Spanish and English words and the development of phonological awareness skills in Spanish-speaking language-minority children. *Reading and Writing*, **29**, 683–704.

Goswami, U. (2010). A psycholinguistic grain size view of reading acquisition across languages. In N. Brunswick, S. McDougall & P. de Mornay Davies (Eds.), *Reading and dyslexia in different orthographies* (pp. 41–60). Psychology Press.

Gottardo, A., & Mueller, J. (2009). Are first-and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, **101**(2), 330.

Hogan, T. P. (2010). A short report: Word-level phonological and lexical characteristics interact to influence phoneme awareness. *Journal of Learning Disabilities*, **43**(4), 346–356.

Hogan, T. P., Bowles, R. P., Catts, H. W., & Storkel, H. L. (2011). The influence of neighborhood density and word frequency on phoneme awareness in 2nd and 4th grades. *Journal of Communication Disorders*, **44**(1), 49–58.

Hogan, T. P., Catts, H. W., & Little, T. D. (2005). The relationship between phonological awareness and reading. *Language, Speech, and Hearing Services in Schools*, **36**(4), 285–293.

Hualde, J. I. (2005). *The sounds of Spanish with audio CD*. Cambridge University Press.

Hualde, J. I., Olarrea, A., Escobar, A. M., Travis, C. E., & Sanz, C. (2020). *Introducción a la lingüística hispánica*. Cambridge University Press.

Jiménez, J. E. (1992). Metaconocimiento fonológico: estudio descriptivo sobre una muestra de niños prelectores en edad preescolar. *Infancia y aprendizaje*, **15**(57), 49–65.

Jiménez González, J. E., & García, C. R. H. (1995). Effects of word linguistic properties on phonological awareness in Spanish children. *Journal of Educational Psychology*, **87**(2), 193.

Jiménez González, J. E., & Ortiz Gonzalez, M. R. (1994). Phonological awareness in learning literacy. *Intellectica*, **18**(1), 155–181.

Kilpatrick, D. A. (2012). Phonological segmentation assessment is not enough: A comparison of three phonological awareness tests with first and second graders. *Canadian Journal of School Psychology*, **27**(2), 150–165.

Lallier, M., & Carreiras, M. (2018). Cross-linguistic transfer in bilinguals reading in two alphabetic orthographies: The grain size accommodation hypothesis. *Psychonomic Bulletin & Review*, **25**(1), 386–401.

Lonigan, C. J. (2006). Conceptualizing phonological processing skills in prereaders. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 77–100). Guilford Press.

Lonigan, C. J., Burgess, S. R., Anthony, J. L., & Barker, T. A. (1998). Development of phonological sensitivity in 2-to 5-year-old children. *Journal of Educational Psychology*, **90**(2), 294.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLOS ONE*, **7**(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Martínez, N., & Goikoetxea, E. (2020). Predictors of reading and spelling words change as a function of syllabic structure in Spanish. *Psicología Educativa. Revista de los Psicólogos de la Educación*, **26**(1), 37–48.

McBride-Chang, C. (1995). What is phonological awareness?. *Journal of Educational Psychology*, **87**(2), 179.

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, **34**(1), 114–135.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, **138**(2), 322.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, **91**(1), 3.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restrucuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Erlbaum.

Míguez-Álvarez, C., Cuevas-Alonso, M., & Saavedra, Á. (2022). Relationships between phonological awareness and reading in Spanish: A meta-analysis. *Language Learning*, **72**(1), 113–157.

Perrachione, T. K., Ghosh, S. S., Ostrovskaya, I., Gabrieli, J. D., & Kovelman, I. (2017). Phonological working memory for words and nonwords in cerebral cortex. *Journal of Speech, Language, and Hearing Research*, **60**(7), 1959–1979.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. The University of Chicago Press.

Roth, F. P., Troia, G. A., Worthington, C. K., & Handy, D. (2006). Promoting awareness of sounds in speech (PASS): The effects of intervention and stimulus characteristics on the blending performance of preschool children with communication impairments. *Learning Disability Quarterly*, **29**(2), 67–88.

Sasisekaran, J., & Byrd, C. (2013). Nonword repetition and phoneme elision skills in school-age children who do and do not stutter. *International Journal of Language & Communication Disorders*, **48**(6), 625–639.

Scarborough, H. S. (2005). Developmental relationships between language and reading: Reconciling a beautiful hypothesis with some ugly facts. In H. W. Catts & A. Kahmi (Eds.), *The connection between language and reading disabilities* (pp. 3–24). Mahwah.

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology*, **91**(3), 439–449.

Seymour, P. H., Aro, M., Erskine, J. M., & Collaboration with COST Action A8 *Network*. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174.

Signorini, A. (1997). Word reading in Spanish: A comparison between skilled and less skilled beginning readers. *Applied Psycholinguistics*, **18**(3), 319–344.

Stage, S. A., & Wagner, R. K. (1992). Development of young children's phonological and orthographic knowledge as revealed by their spellings. *Developmental Psychology*, **28**(2), 287.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology*, **38**(2), 175–190.

StataCorp. (2023). Stata statistical software: Release 17. StataCorp LLC.

Swan, D., & Goswami, U. (1997). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of Experimental Child Psychology*, **66**(1), 18–41.

Swank, L. K., & Catts, H. W. (1994). Phonological awareness and written word decoding. *Language, Speech, and Hearing Services in Schools*, **25**(1), 9–14.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **58**(1), 267–288.

Treiman, R., & Weatherston, S. (1992). Effects of linguistic structure on children's ability to isolate initial consonants. *Journal of Educational Psychology*, **84**(2), 174.

Treiman, R., & Zukowski, A. (1991). Levels of phonological awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 67–83). Lawrence Erlbaum Associates.

Troia, G. A., Roth, F. P., & Yeni-Komshian, G. H. (1996). Word frequency and age effects in normally developing children's phonological processing. *Journal of Speech, Language, and Hearing Research*, **39**(5), 1099–1108.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades?. *Journal of Child Psychology and Psychiatry*, **45**(1), 2–40.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, **101**(2), 192.

Wolf, M., O'rourke, A. G., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing*, **15**, 43–72.

Wydell, T. N., & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition*, **70**(3), 273–305.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, **101**(476), 1418–1429. https://doi.org/10.1198/016214506000000735

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, **131**(1), 3.

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press

Appendix A
Measure development, calibration, and validation in English

Measure development

The initial word list was based on curricula from California Dual Language Programs and later expanded by researchers to meet selection criteria for Elision Receptive and Expressive tasks. Words were chosen such that removing a linguistic unit (word, syllable, or phoneme) produced a new word. Deletions occurred at the beginning or end (K–G2) or in the middle (G2 only). Items with repeated syllables or phonemes (e.g., blends with [ɹ] or /l/) were excluded to ensure phonetic diversity.

Receptive tasks required selecting an image representing the target word, emphasizing high imaginability and cultural neutrality. Expressive tasks involved repeating the word without the deleted segment and included real and nonwords, especially for G1–G2 students.

Participants

See Table S1 in Supplementary Materials.

Procedures

Students completed the Elision task using yoked iPads, with audio prompts and untimed administration. Receptive tasks presented three images, and students selected the correct one. Testing took place in quiet school settings, and responses were scored dichotomously.

Psychometric properties

Items were calibrated using a Rasch model (TAM package in R). We checked unidimensionality and removed poorly fitting items (MSE > 2). Receptive and expressive elision items were jointly calibrated because the goal of the study was to model item difficulty along a common phonological awareness continuum. Consistent with a developmental view of phonological awareness as a unitary construct with heterotypic continuity (Anthony & Francis, 2005; Anthony et al., 2011; Schatschneider et al., 1999), receptive and expressive formats were treated as different task manifestations of the same underlying ability. Differences between formats were therefore modeled as systematic sources of variation in item difficulty rather than as indicators of separate constructs, allowing all items to be placed on a common scale. The final pool included 157 English items and 172 Spanish items. Person separation reliability was high: 0.823 (English), 0.801 (Spanish GK–1), and 0.748 (Spanish G2).

To assess criterion validity, we examined correlations between the Multitudes elision tasks and established standardized measures. The elision expressive task in English showed a moderate correlation with the CTOPP-2 elision subtest ($r = 0.66$), while the elision receptive task had a slightly lower but still moderate correlation ($r = 0.48$). Similarly, with the TPAS deletion subtest, the elision expressive and receptive tasks showed correlations of 0.45 and 0.43, respectively. These findings provide evidence of moderate criterion validity. As expected, receptive task correlations were slightly lower due to differences in response format; both CTOPP-2 and TPAS rely on expressive responses, which may align more closely with the Multitudes elision expressive format.

Appendix B

See Tables S2 and S3 in Supplementary Materials.

Cite this article: Zegers, M., Siebert, J. M., Durán, L. K., Jasso, J., Pratt, A. S., Lichtmacher, J., Murphy, H., Pei, F., & Gorno-Tempini, M. L. (2026). Features predicting item difficulty in English and Spanish elision task. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716426100976

https://doi.org/10.1017/S0142716426100976 Published online by Cambridge University Press