

Between Grammatical Rigidity and Pragmatic Flexibility: Viewing the Syntax-Prosody-Pragmatics Interface Interaction through an “a + Pause” Insertion Test

Wei Shi

2026-07-17

ChinaRxiv

View the original and related papers at
<https://chinaxiv.org/items/chinaxiv-202607.00169>

Source: ChinaXiv. Machine translation. Verify with the original.

Abstract

Based on the Chinese sequence “we should study hard,” this paper constructs a set of minimal pairs to examine the gradient of acceptability for “a + pause” in different syntactic positions: a pause after the subject is natural, a pause after a functional head is awkward, and a pause within a prosodic word is by default illicit. This sequence reveals a theoretical paradox: why are illicit pauses within prosodic words tolerated and successfully interpreted by native speakers in special pragmatic contexts such as “earnest exhortation” or “stuttering”? This paper successively evaluates the explanatory power of formal syntax, prosodic phonology, and interactional linguistics, arguing that formal approaches can precisely predict boundary strength but cannot accommodate pragmatic counterexamples, whereas functional/interactional approaches can explain contextual tolerance but have difficulty predicting the hierarchy of structural rigidity. A small-scale pilot experiment with 39 participants preliminarily verifies the group-level robustness of the boundary-strength gradient and the pragmatic override effect. On this basis, this paper proposes a “core-periphery dual-track interaction model,” analyzing the language-processing system as the coordinated operation of a “core algorithmic track” for rigid structural computation and a “peripheral pragmatic-inference track” for contextualized intention reasoning and fault-tolerant exemption. “A + pause” is precisely a minimal window for observing the switching and negotiation between the two tracks.

Full Text

Between Grammatical Rigidity and Pragmatic Flexibility: Viewing the Syntax-Prosody-Pragmatics Interface Interaction through an “a + Pause” Insertion Test

Wei Shi

(School of Foreign Languages, Jingdezhen University, Jingdezhen 333400)

Abstract: Using the Chinese sequence “我们要好好学习” (“we should study hard”) as the basis, this paper constructs a set of minimal contrasting sentences to examine the acceptability gradient of “a + pause” across different syntactic positions: a pause after the subject is natural; a pause after a functional head is awkward; and a pause within a prosodic word is by default illicit. This sequence reveals a theoretical paradox: why, in special pragmatic contexts such as earnest, emphatic speech or stuttering, can an otherwise illicit pause within a prosodic word be tolerated and successfully interpreted by native speakers? This paper evaluates, in turn, the explanatory power of formal syntax, prosodic phonology, and interactional linguistics. It argues that formal approaches can precisely predict boundary strength but cannot accommodate pragmatic counterexamples,

whereas functional/interactional approaches can explain contextual tolerance but have difficulty predicting the hierarchy of structural rigidity. A small-scale pilot experiment with 39 participants preliminarily confirms the group-level robustness of the boundary-strength gradient and the pragmatic override effect. On this basis, the paper proposes a “Core-Periphery Dual-Track Interaction Model,” analyzing the language-processing system as the coordinated operation of a “Core Algorithm Track” for rigid structural computation and a “Peripheral Pragmatic Inference Track” for contextualized intention reasoning and fault-tolerant exemption. The “a + pause” construction thus provides a minimal window for observing switching and negotiation between the two tracks.

Key words: sentence-final/medial particle “a” (啊); pause; syntax-prosody interface; prosodic word; dual-track model; pragmatic override

Classification number: H146.3

Between Grammatical Rigidity and Pragmatic Flexibility: Insights into the Syntax-Prosody-Pragmatics Interface from the Chinese “a + Pause” Insertion Test

Wei SHI (E-mail: 20031@jdzu.edu.cn)

(School of Foreign Languages, Jingdezhen University, Jingdezhen 333400, China)

Abstract: Using a minimal paradigm based on the Chinese sequence *wǒmen yào hǎohǎo xuéxí* (我们要好好学习), this study investigates the acceptability gradient of “a + pause” sequences across three syntactic positions: post-subject (natural), post-functional-head (awkward), and inside a prosodic word (default-ungrammatical). This robust hierarchy gives rise to a core theoretical paradox: prosodic-word-internal pauses, normally ruled out as illicit in neutral contexts, become tolerable and interpretable for native speakers under specific pragmatic conditions such as emphatic rhetoric or disfluent speech. To resolve this paradox, the paper assesses the explanatory limits of three dominant linguistic frameworks—formal syntax, prosodic phonology, and functional-interactional linguistics. It shows that formal approaches precisely predict the hierarchical strength of syntactic-prosodic boundaries but cannot account for pragmatic tolerance of violations, whereas functional-interactional approaches explain contextual malleability but fail to derive the structural rigidity gradient. A pilot acceptability judgment study with 39 native speakers provides preliminary empirical support for the predicted hierarchy and the pragmatic override effect. Building on these findings, we propose the Core-Periphery Dual-Track Interaction Model, which posits that language processing relies on the coordinated operation of an automatic, modular Core Algorithm Track (for structural computation) and a context-sensitive Peripheral Pragmatic Inference Track (for intention reasoning

and fault-tolerant exemption). The “a + pause” construction thus serves as a minimal empirical window for observing the dynamic switching and negotiation between the two tracks.

Key words: sentence-medial particle *a* (啊); pause; syntax-prosody interface; prosodic word; core-periphery dual-track model; pragmatic override

1. Introduction: A Theoretical Puzzle Raised by a Minimal Contrast Set

Please compare the acceptability of the following four sentences in a default neutral context :

- (1) “我们要好好学习。” (“We must study hard.”)
(baseline: no disfluency; completely natural)
- (2) “我们啊, 要好好学习。” (“We, ah, must study hard.”)
(pause after the subject: natural; topicalized)
- (3) “我们要啊, 好好学习。” (“We must, ah, study hard.”)
(pause after a functional head: awkward; highly marked)
- (4) “* 我们要好啊, 好学习。”
(pause inside a prosodic word: unacceptable by default)

These four sentences form a clean set of minimal matched sentences. The only variables are the modal particle “啊” and the position at which the accompanying pause is inserted. Their acceptability displays a clear gradient: from the complete naturalness of sentence (1) , through the naturalness of sentence (2) , and the awkwardness of (3) , to the default unacceptability of sentence (4) . This gradient intuitively reflects speakers’ tacit knowledge of the degree of internal tightness in syntactic structure, and it has also been group-validated in a small-scale pretest with native speakers (see §2.4) . However, what constitutes the real challenge is the “dual identity” of sentence (4) .

In a default context, sentence (4) is consistently judged by native speakers to be a “wrong sentence.” This judgment is so stable that it can be made without any linguistic training. Yet once the context is set to either of the following two situations, sentence (4) can be tolerated by native speakers, and may even be regarded as an acceptable expression:

- (5) Context A: A teacher earnestly says to students — “我们要好啊, 好学习! 不能再贪玩了。”
“We must do well—study hard! We cannot keep fooling around.”
- (6) Context B: Because of nervousness, the speaker produces an unexpected disfluency in speech — “我们要好啊……好学习。”
“We must do well, ah…study hard.”

This gives rise to what this paper calls the “Sentence 4 Paradox”: the same linear sequence is judged by core syntax to be extremely unacceptable in a default context, yet in a specific pragmatic context it can acquire marginal interpretability through the limiting permeation of the periphery. The paradox involved here is far more than an isolated case. It points directly to a fundamental issue in linguistic theory: when we call a sentence “legal” or “illegal,” are we ultimately discussing a syntactic concept or a pragmatic concept? How can the “rigidity” of syntactic rules and the “flexibility” of contextual effects coexist within the same cognitive system?

Before answering these questions, it is first necessary to clarify three concepts that are often conflated but that, in this paper, must be strictly distinguished. First, grammaticality. Grammaticality is determined by core syntactic rules. Sentence (4) is ungrammatical in any context: a pause has entered the interior of a prosodic word and has triggered the highest-level violation signal of the core track. The computational output of this core track is not erased no matter how the context changes. Second, acceptability. Acceptability is the level at which native speakers make psycholinguistic judgments. Sentence (4) is unacceptable in a default context; but under the pragmatic interventions in (5)–(6), it can obtain a kind of “marked acceptability” —native speakers will not take it to be “a mistake,” but will instead regard it as “said deliberately in this way” or as “getting stuck.” Third, intelligibility. Intelligibility is the decoding of information at the interactional level. Under pragmatic intervention,

The core phenomenon discussed in this paper has been group-validated through a small-scale pretest with 39 participants (see §2.5 for details). In the pretest, the mean acceptability score for sentence(3) was 2.82. There was some individual variation, but most participants reported it as “awkward” or “not very natural,” rather than as an “absolute error” on a par with sentence(4). The “*” prefixed to sentence (4) indicates only its illegality in a default neutral context, and does not apply to marked contexts involving pragmatic intervention. Detailed statistical distribution data from the pretest are provided in the supplementary materials.

of sentence (4) remains intelligible throughout, and the hearer can accurately recover the propositional content the speaker intended to express: “we must study hard.”

Now the core claim of this paper can be stated in advance: the “pragmatic rescue” or “pragmatic covering” discussed below in no way changes the grammaticality of the altered sentence (4)—it is always grammatically deviant. Rather, it shifts the source of its acceptability from the system of syntactic rules to the system of pragmatic inference. In other words, the hearer no longer evaluates this sequence as a “syntactic product,” but reframes it as a “product of rhetorical strategy” or as a “linguistic symptom.” This operation of “reframing” is precisely the core mechanism that we need to model.

Taking the “sentence (4) paradox” as its theoretical touchstone, this paper

attempts to test, in turn, the explanatory power of three major linguistic approaches, to demonstrate the unavoidable “explanatory blind spots” each contains, and on that basis to propose an integrative theoretical model—the core-periphery dual-track interaction model. The line of argument is as follows: Section 2 constructs a diagnostic tool, proposing the operational logic and theoretical assumptions of the “ah + pause” insertion test; Section 3 develops the theoretical confrontation, using the “sentence (4) paradox” to interrogate both the formal path and the functional/interactional path; Section 4 proposes and elaborates the dual-track model, distinguishing two pragmatic-covering mechanisms of different natures; Section 5 summarizes the article and points out its limitations and future directions.

2. The “ah + pause” insertion test: a diagnostic tool for the strength of syntactic-prosodic boundaries

2.1 Theoretical predictions for boundary-strength levels

In the sequence “we must study hard” discussed in this paper, “must” is a modal verb; within the generative-syntactic framework it belongs to the functional category T (the head), and its projectional structure is:

$$(7) [TP \text{ we } [T' \text{ must } [vP \text{ study hard }]]]$$

This is the syntactic starting point for all subsequent discussion: the verb phrase “study hard” following “must” is the complement of the head T, “must.” The two form a “head-complement” structural relation, one of the tightest sisterhood relations in the syntactic tree.

On the basis of the basic principle of structuralist immediate-constituent analysis—that pauses can occur only at the boundaries of immediate constituents—and the constraints in generative grammar concerning the head-complement relation, we can predict in advance three levels of “boundary permeability” :

- (8) Boundary-permeability hierarchy: maximal-projection boundary (after the subject) > head-complement boundary (between T and vP) > prosodic-word-internal boundary

The first level is the maximal-projection boundary, between the subject “we” and T'. This is the highest-level boundary within TP and is fully open to the insertion of pauses and topic markers; the naturalness of sentence (2) is thereby predicted. The second level is the head-complement boundary, between “must” and “study hard.” Because of the strict government relation between the head and its complement, this position can receive only a low-level prosodic boundary in the syntax-prosody mapping. It cannot satisfy the requirement that “ah + pause” attach to a phonological-phrase boundary, and therefore it strongly resists the insertion of pauses and independent elements; the awkwardness of sentence (3) is thereby predicted. The third level is the prosodic-word-internal boundary, between “hard” and “hard” in the reduplicated form. This is the

lowest-level boundary and by default is extremely resistant to penetration; the unacceptability of sentence (4) is thereby predicted.

2.2 Boundary basis: the indivisibility of the prosodic word

There is a potential controversy in the discussion that follows. Is the “hard-hard” in “study hard-hard” really a “word” in the strict sense? If it is, then what sentence (4) violates is the Lexical Integrity Hypothesis; but if there is a division

...if the analysis holds that “好好” is an independent adverbial modifier, then what sentence (4) violates is not a lexical boundary, but something else.

This paper therefore chooses to ground its argument in the more robust evidence of prosodic phonology, so as to bypass the controversy over the morphological status of “好好” : regardless of whether “好好” is analyzed syntactically as a reduplicated adverb or as a compound, at the prosodic level it constitutes a standard disyllabic prosodic word—a foot (Feng Shengli, 1997). “学习” likewise is an independent disyllabic prosodic word; when the two are combined, “好好学习” forms a prosodic phrase at a higher level (Duanmu San, 2000). Within the general framework of prosodic hierarchy theory (Selkirk, 1984; Nespor & Vogel, 1986), the prosodic word is the smallest independent unit in the prosodic hierarchy, and its internal structure is by default impermeable—under default conditions it strongly resists the insertion of any kind of pause, modal particle, or other extra-prosodic element.

Thus, the unacceptability of sentence (4) lies in the fact that the modal particle “啊,” together with its accompanying pause, attempts to insert into the internal morpheme boundary of the disyllabic prosodic word “好好” ; under default conditions, this position displays an extreme resistance to the insertion of any independent element.

Of course, the impermeability of the interior of a prosodic word and the Lexical Integrity Hypothesis overlap to a high degree in their empirical outcomes, but their theoretical foundations are nevertheless different: the former is based on prosodic-hierarchical mapping, whereas the latter is based on the modular division between the lexicon and syntax.

2.3 The Complete Logic of the Probe

To elevate the insertion of “啊 + pause” into a methodological tool, it is necessary to clarify how it works. This paper proposes the following logical chain:

depth of syntactic hierarchy → rank of the mapped prosodic boundary → whether it can host “啊 + pause”
(9)

Prosodically, the modal particle “啊” and its accompanying pause must normally attach to the right boundary of a prosodic unit whose minimum rank is that of

a “phonological phrase” ; they cannot attach to a prosodic word or to a still smaller unit. Lü Shuxiang (1956) observed relatively early the pausal use of the modal particle “啊” in Chinese sentences; all of his examples occur after larger syntactic constituents such as subjects and topics, providing early distributional evidence for subsequent research on boundary hierarchy. Xiong Ziyu and Lin Maocan (2004), through acoustic experiments, verified that the boundary at which sentence-internal “啊” occurs is consistently accompanied by significant syllable lengthening and pitch resetting, forming a sharp contrast with the absence of a perceptible pause inside a prosodic word. Wang Wei and Bi Yanjuan (2017) further clarified, from the perspective of formal classification, that the legitimacy of sentence-internal modal particles is positively correlated with the syntactic-prosodic level of the constituent to which they attach: the higher the level, the stronger the acceptability.

After “啊 + pause,” as a clitic, attaches to a prosodic host, it may form a local clitic group at the PF level. However, this PF-level prosodic restructuring does not alter the boundary-strength rank output by the core track at the stage of syntax-prosody mapping: the T-vP boundary has already been marked in the core-track computation as “strongly resistant to penetration” that is, the syntactic government relation is too tight, and the rank of the projected prosodic boundary is insufficient to host the insertion of an independent element. The attachment of “啊 + pause” is a matter of PF-level implementation detail, not a recalculation of syntax-prosody mapping. Thus, “the boundary rank output by the core track” and “PF-level prosodic realization” are distinguished as two separate levels.

This means that the acceptability gradient of the four sentences is not sometimes attributable to “syntax” and sometimes to “prosody.” Rather, it uniformly reflects one fact: the deeper the syntactic structural level and the tighter the syntactic relation, the lower the rank of the prosodic boundary it projects, and the weaker its capacity to host “啊 + pause.” This is the theoretical basis for using the “啊 + pause” insertion test as a probe of syntactic-prosodic boundary strength.

2.4 Preliminary Verification in a Pre-experiment

To test whether the above gradient of boundary strength is a general intuition among native speakers rather than the author’s personal introspection, this paper conducted a small-scale acceptability pretest to preliminarily verify the group robustness of the “a + pause” insertion test.

The participants were 39 native speakers of Mandarin Chinese, all second-year Business English majors, aged 18-20, with no systematic training in linguistics (none had systematically studied courses related to modern Chinese or linguistics); interference from linguistic knowledge was thus excluded.

The experiment used a 7-point Likert scale (1 = completely disfluent, 7 = completely natural). The test materials included 4 core test sentences and 12 filler sentences, all presented in random order. The participants were guided by the

instructions to activate spoken-language representations through silent reading, to make rapid judgments based on first impressions, and to avoid grammatical deliberation. The mean acceptability scores for the core test sentences in a neutral context are shown in Table 1:

Table 1: Mean acceptability scores for core test sentences in a neutral context

Syntactic position	Example sentence	Mean acceptability	Theoretical prediction
Baseline (no insertion)	We must study hard	6.44	Completely natural
After the subject (maximal projection boundary)	We, <i>ah</i> , must study hard	5.67	Natural; highly permeable
After a functional head (head-complement boundary)	We must <i>ah</i> , study hard	2.82	Awkward; strongly resists penetration
Within a prosodic word	We must study ha- <i>ah</i> , hard	1.90	Default: unacceptable; extremely resistant

The results show that acceptability across the four positions presents a significant decreasing gradient, fully matching the theoretically predicted hierarchy of boundary permeability. In particular, the mean score for the position within a prosodic word was only 1.90, and 66.7% of the participants directly assigned the lowest score, indicating that the impenetrability of prosodic words in a neutral context is a strong consensus among native speakers.

The generalization test further showed that, for three groups of sentences in which “*ah*” was inserted after other modal verbs— “We have to *ah*, take this matter seriously,” “He can *ah*, run five kilometers in one breath,” and “They will *ah*, complete the task on time”—the mean acceptability scores were concentrated in the range of 2.92–3.33, highly overlapping with the score of 2.82 for “We must *ah*, study hard.” This indicates that the strong resistance to particle insertion after a functional head is not an isolated case involving “must,” but a general pattern of functional categories in Chinese; the “*ah* + pause” insertion test, as a diagnostic tool for boundary strength, has cross-lexical generality.

In sum, the pretest results provide preliminary empirical support for the group robustness of the core phenomenon and verify the mapping logic of “depth of

syntactic hierarchy $\rightarrow$ level of prosodic boundary $\rightarrow$ acceptability.” The theoretical discussion below is built on the basis of this group-validated phenomenon. The complete experimental materials, instructions, and contextualized testing scenarios are provided in the Appendix.

3. Theoretical Confrontation: Testing Explanatory Power with the “Sentence 4 Paradox”

With the diagnostic tool and theoretical predictions in place, we can now enter the core theoretical confrontation. The aim of this section is not to provide a neutral summary of the various schools, but to use the “Sentence 4 paradox” as the same benchmark and simultaneously test two

the predictive power and explanatory boundaries of the main approaches.

3.1 Precision and Boundaries of the Formal Path

The formal path—including structuralist immediate-constituent analysis, the head-government theory of generative grammar, and the prosodic-hierarchy theory of prosodic phonology—performs exceptionally well in predicting gradients of acceptability in default contexts.

For sentence (2), the formal path correctly predicts that the boundary after the subject is a natural pause position. For sentence (3), it correctly predicts the strong rejection of a pause between the head and its complement: the head must strictly govern its complement, and the insertion of a pause and a modal particle destroys this government relation. For sentence (4), it correctly predicts the near-impenetrability of the interior of a prosodic word.

Up to this point, the formal path has made no errors. Yet when we present it with the variants (5)–(6) of sentence (4), it can offer only one answer: “This belongs to phenomena at the level of language use and does not fall within the scope of core syntactic government.” This answer is not itself wrong—indeed, we consider it correct. But it exposes a “theoretical boundary”: the formal path delineates the precise contours of the “rigid core,” yet it cannot—and has no intention to—provide a mechanistic account of how the core interacts with peripheral pragmatic modules. If we acknowledge that these two phenomena—syntactic judgment and pragmatic tolerance—coexist in the cognitive system of native speakers, then a complete linguistic theory should account for the interface between them, rather than merely declaring that “this is the business of another department.”

3.2 Tolerance and Limits of the Functional/Interactional Path

On this issue, the functional/interactional path displays precisely the opposite strengths and weaknesses. Faced with the “revival” of sentence (4) in (5)–(6), interactional linguistics and relevance theory can provide convincing explanations. For (5), the analytic path of Sperber and Wilson (1995) is as follows: the

hearer recognizes the speaker's higher-order communicative intention of being "earnest and emphatic," reinterpreting the forceful dismantling of the prosodic word in "hǎo-ā-hǎo" as an intensifying rhetorical strategy, thereby shifting the source of acceptability from "syntactic conformity" to "rhetorical effectiveness." For (6), the analysis of interactional linguistics by Couper-Kuhlen and Selting (2018) is that stuttering is a typical "self-initiated-self-repair" sequence; the hearer automatically identifies it as an obstacle in speech production and thus suspends the judgment of syntactic legality.

However, the functional/interactional path faces a thorny problem: if the same context of "earnest emphasis" can reduce the awkwardness of sentence (3), why can it only make sentence (4) barely acceptable, rather than bring it close to naturalness? More sharply put: why is the upper limit of naturalness after contextual accommodation strictly negatively correlated with the boundary strength of the core axis—the more rigid the boundary, the lower the "ceiling of naturalness" after accommodation?

This shows that pragmatic tolerance is not a pass with "unlimited quota." It always encounters a rigid resistance originating in the deep structure. The functional/interactional path can explain "why tolerance is possible," but it cannot predict "the gradient of difficulty in tolerance." Yet precisely this gradient has already been accurately predicted by the formal path on the basis of syntactic hierarchical depth.

3.3 Interim Summary at the Intersection

The two paths each have their insights, and each has its blind spots: the formal path can perfectly answer "why it is wrong," but cannot answer "why it is still usable despite being wrong"; the functional/interactional path can perfectly answer "why it is still usable despite being wrong," but cannot answer "why errors in different positions differ in how difficult they are to use."

This is not a simple defect of either theory, but rather the result of the "design choices" inherent in their respective theoretical architectures.

The architecture of the formal pathway makes it adept at handling "default," "automatic," and "modular" computation, whereas the architecture of the functional/interactive pathway makes it adept at reasoning about "context," "intention," and "interaction." The question is: how can these two systems coexist in the brain of a native speaker and operate collaboratively in real-time language processing? This is the true theoretical task left to us by the "Sentence 4 Paradox."

4. Thesis: The "Core-Periphery" Dual-Track Interaction Model

On the basis of the theoretical confrontation in Section 3, we attempt to propose an integrative framework. The core idea of this framework is not to regard

the formal pathway and the functional/interactive pathway as “two competing explanations of the same phenomenon,” but rather to redefine them as “the operations of two distinct cognitive subsystems.” The two exist in parallel, perform their respective functions, and may interact under specific conditions.

4.1 Model Architecture

We propose the “Core-Periphery Dual-Track Interaction Model.” Its basic architecture comprises two subsystems: the “core algorithmic track” and the “peripheral pragmatic-inferential track.”

The nature of the “core algorithmic track” is automatic, mandatory, and modular. Its function is to carry out the hierarchical computation of syntactic structure and the mapped output of prosodic boundaries. More specifically, the core track can be subdivided into two submodules: the syntactic computation module is responsible for generating hierarchical structure, while the PF prosodic mapping module is responsible for converting syntactic structure into levels of prosodic boundary strength. Its output takes the form of a continuous value of structural boundary strength, ranging from “fully permeable” (e.g., the post-subject boundary), to “strongly resistant to permeation” (the head-complement boundary), to “extremely resistant to permeation” (the boundary within a prosodic word). Its operational characteristic is that it remains online by default, is not affected by context or intention, and its computational result constitutes the default basis of acceptability.

The nature of the “peripheral pragmatic-inferential track” is context-sensitive, intention-driven, and based on the principle of relevance. Its function is to monitor the communicative situation, identify the speaker’s intention, and process special pragmatic signals (such as rhetorical strategies and speech impediments). Its operational characteristic is that, in a default neutral context, it remains “silent” or in a low-activation state; when specific communicative cues are detected, it is activated and can initiate an operation of “identity reframing.”

With regard to the real-time coordination of the two tracks, this model adopts a hybrid position of “default seriality and conditional parallelism”: the structural computation of the core track is always online and is not affected by context; contextual information from the peripheral track continuously permeates in the form of “top-down expectations,” but in a default neutral context it remains in a low-activation state and does not interfere with the output of the core track. Only when the core track outputs a “violation” signal does the peripheral track trigger an explicit operation of “identity reframing.” This preserves the core predictive power of modularity while also accommodating the context-sensitivity of interaction.

Finally, this model needs to be distinguished from several existing dual-system theories. Unlike the static “competence-performance” distinction, this model adopts a dynamic processing perspective and focuses on the interaction mechanism between the two systems in real-time processing. It is compatible with

the “coding-inference” framework of Relevance Theory proposed by Sperber and Wilson (1995), but further refines the mechanism by which the inferential system performs “identity reframing” on syntactic output. It is also compatible with findings in neurocognitive research on language processing concerning the early automaticity of syntactic processing, as in Friederici (2002). This model is consistent in theoretical orientation with Jackendoff’s (1997, 2002) Parallel Architecture, in that both maintain that multiple systems interact in parallel. However, it focuses on the hierarchical division between the “core language system” and the “peripheral pragmatic system,” rather than on parallelism among different representational levels within language; the two are situated at different dimensions of analysis.

It should be added that this article focuses on the specific interface phenomenon of pause distribution. The core track adopts the unidirectional mapping from “syntactic structure $\rightarrow$ prosodic boundary level” as a working hypothesis, with the aim of revealing the foundational constraining role of syntactic hierarchy on prosodic boundaries. The reverse shaping effect of prosodic rules on syntactic structure belongs to syntax-prosody interface research

another independent dimension, which lies outside the core scope of the present discussion and will be left for future work.

4.2 Synergistic Mechanism: Default Mode and Intervention Mode

The synergistic operation of the dual tracks can be divided into two modes.

Mode 1: Default mode (core-track dominance). In a neutral context, the peripheral track remains silent. The acceptability judgment of the utterance is determined entirely by the strength of the structural boundary in the core track. The acceptability of the four sentences directly maps onto the hierarchy of boundary permeability:

$$\text{Sentences 1, 2} \gg \text{Sentence 3} \gg \gg \text{Sentence 4} \quad (10)$$

This is a clear and predictable gradient. At this point, no intervention by “pragmatic override” is needed.

Mode 2: Intervention mode (peripheral-track override). When specific “override-triggering signals” are present in the context, the peripheral track is activated and performs an override operation. The core mechanism of override is to “reframe the identity” of the “violating structure” output by the core track.

One might ask whether the hearer could temporarily reanalyze the first *hao* “good” in example (4) as an independent response word, thereby shifting the insertion site of *a* from within the prosodic word to a legitimate boundary. The answer is no. From the perspective of syntactic position, the response word *hao* occurs only sentence-initially or at the left edge of a projection, whereas the

first *hao* in example (4) is in an internal adverbial position within the vP after a modal verb; the categorial difference is substantial, and there is no corresponding syntactic movement mechanism to support such a shift. From the perspective of semantic autonomy, if the first *hao* were segmented as an independent word, the subsequent *hao xuexi* “study well” could not form a legitimate syntactic-semantic combination, and the reanalysis would create a new conflict.

The preliminary experimental results provide corroborating evidence for this: if a reanalysis route were available during processing, the acceptability of sentence (4) in a neutral context should not be significantly lower than that of sentence (3). However, the observed results show that the mean for sentence (4) is only 1.90, far below the 2.82 for sentence (3), with an ordinal difference approaching 1 point. This indicates that native speakers, by default, cannot resolve the violation through reanalysis; the core track’s boundary judgment is processing-mandatory.

Therefore, the structural boundary output by the core track remains, by default, non-adjustable during processing. The peripheral track does not modify the computational result of the core track—the grammaticality judgment remains unchanged—but instead reframes the “violating structure”: it is no longer treated as a “linguistic structural product” requiring syntactic evaluation, but is assigned a new communicative identity.

Peripheral-track override is not a single type. According to the nature of the triggering signal and the strength of exemption, it can be divided into two subtypes. What the two share is that neither changes the grammaticality judgment output by the core track; they differ in the logic by which the violation is attributed and in the degree of exemption.

First, strategic override (production side). The triggering signal is an intentional rhetorical purpose, such as the earnest emphatic posture in example (5). The speaker actively decomposes the prosodic word, using *a* and the pause as devices for rhetorical reinforcement; the hearer reframes the utterance as a “product of rhetorical strategy,” and the acceptability judgment shifts from “whether the syntax is correct” to “whether the rhetoric is effective.” This is an active violation operation led by the speaker. Because the underlying violation signal is still retained, the strength of exemption is relatively limited. The preliminary experimental data show that this type of override increases the acceptability of a pause inside a prosodic word by 0.77 points, lower than that of tolerance-based override.

Second, tolerance-based override (comprehension side). The triggering signal is an unexpected pause; typical cases include production impairments such as stuttering and online disfluencies caused by interruptions in thought. The hearer attributes the formal anomaly to a “failure of the non-linguistic system” or to “online processing pressure,” thereby exempting the utterance as a whole from syntactic well-formedness judgment. This is a passive exemption operation led by the hearer. Because the source of the violation is moved outside the linguistic

system itself, the strength of exemption is higher. The preliminary experimental data show

shows that the increase in acceptability produced by this type of coverage for pauses inside a prosodic word is 1.07 points, significantly higher than that of strategic coverage.

The difference in strength between the two types of coverage further corroborates this point: the violation signal on the core track always remains and is not erased by pragmatic intervention; the strength of the coverage effect depends on whether the peripheral track can attribute the violation to something outside the linguistic system, thereby lowering the weight assigned to the core-track output. In error-tolerant coverage, the violation is attributed to a system-external malfunction, and thus the degree of exemption is higher; in strategic coverage, the violation still belongs to an active choice within the linguistic system, and therefore the underlying resistance is stronger.

Within error-tolerant coverage there are also two subtypes. First, there is pathological tolerance, such as physiological failures in speech production including stuttering and articulation disorders; the peripheral track attributes these to “non-linguistic-system malfunctions” and exempts them from syntactic judgment. Second, there is hesitation-type tolerance, such as non-pathological involuntary pauses (for instance, unexpected gaps caused by interruptions in thought or lexical search); the peripheral track identifies these as “involuntary interruptions in the process of online discourse organization,” likewise exempting them from syntactic judgment, but attributing them to “online organizational pressure within the linguistic system” rather than to an “external system malfunction.” What the two subtypes share is that the “source of the violation is attributed to an external cause” ; they differ in the nature of that external cause.

Seen in this way, the sources and nature of the acceptability of (3) and (4) are entirely different: the “awkwardness” of sentence (3) belongs to marked acceptability at the level of the syntactic product—the structure itself is legal, only unnatural; whereas the “marked acceptability” of sentence (4) after pragmatic coverage belongs to rhetorical/error-tolerant recognition at the communicative level—the structure itself is illegal, but is reframed as a rhetorical product or a speech symptom.

4.3 A Unified Interpretation of the Full Dataset by the Model

We can now use the dual-track model to provide a unified interpretation of the performance of all test sentences in different contexts. Table 2 simultaneously lists the theoretical predictions and the measured means from the pre-experiment, showing a high degree of consistency between the two:

Table 2: Comparison between the Interpretive Framework of the Dual-Track Model and the Pre-Experiment Measurements

Example sentence	Context	Core-track output	Peripheral-track state	Final acceptability	Measured mean
Sentence 2 (after the subject)	Neutral	Boundary open	Silent	Unmarked naturalness	5.67
Sentence 3 (after a function word)	Neutral	Strong resistance	Silent	Awkward, highly marked	2.82
Sentence 4 (inside a prosodic word)	Neutral	Extreme resistance	Silent	Unacceptable by default	1.90
Sentence 3 (after a function word)	Earnest and emphatic	Strong resistance	Strategic coverage (awkwardness weakened)	Marked acceptability, tending toward naturalness	3.90
Sentence 4 (inside a prosodic word)	Earnest and emphatic	Extreme resistance	Strategic coverage	Marked acceptability (effortful but effective)	2.67
Sentence 4 (inside a prosodic word)	Stuttering/getting stuck	Extreme resistance	Error-tolerant coverage	Tolerable and comprehensible	2.97

The core implication of this table is that the computational results of the core track are, in processing, by default non-adjustable: the “extreme resistance” property of boundaries internal to prosodic words remains constant across all contexts. Pragmatic intervention changes only the dimension on which acceptability is evaluated, but does not eliminate the underlying signal of structural violation. The pre-experimental data show that, in both pragmatic contexts, the increase in acceptability is negatively correlated with boundary strength: the increase after a function word (about 1.1 points) is higher than that within a prosodic word (about 0.8-1.1 points), and the interior of a prosodic word never breaks through the three-point level of “somewhat awkward.” This precisely confirms the model’s prediction that “the more rigid the boundary, the greater the cost of coverage.”

4.4 Model Prediction: The Limits of Coverage

Finally, the present model can make a prediction about the “limits of coverage.” Before doing so, the dual-track model must first answer a key question head-on: why is the difficulty of pragmatic coverage positively correlated with the boundary strength of the core track? The answer is that the covering operation of the peripheral track is essentially a form of “reverse decoding” : it must use the original cues remaining in the violated structure to recover the speaker’s communicative intention. The more rigid the boundary is—that is, the more the output of the core track resists penetration—the fewer original cues retained in the violated structure, the higher the redundancy of contextual information required for reverse decoding, and thus the greater the pragmatic cost of coverage. This mechanism maps precisely onto all the data: sentence (2) requires no coverage, because the boundary output by the core track is itself open; sentence (3) requires weak contextual triggering, because although the boundary between the central word and the complement is strongly resistant, the structural cues of the whole vP remain intact, so the cost of reverse decoding is moderate; sentence (4) requires activation by an extremely strong context, because resistance is highest at the boundary internal to the prosodic word, and only the two neighboring traces of “good” and the combinatorial cue with “learn” remain available for reverse use. This gradient is itself the core empirical evidence for the dual-track model, rather than a residual phenomenon awaiting explanation.

Where, then, are the limits of coverage? Not every syntactic violation can be rescued by pragmatic coverage. The condition for successful coverage is that the “violated sequence” output by the core track still preserves enough internal structural cues for the peripheral track to “reverse-decode” the original communicative intention.

The reverse decoding of sentence (4) relies on the semantically independent redundant residue of “good-good” : the adjacency repetition of the two instances of “good,” together with their combinatorial relation with the following “learn,” jointly activates the overall lexical pathway. This mechanism can be extended to cases that are more difficult to cover. Non-overlapping disyllabic words (for example, “serious”) can also be penetrated in extreme contexts (for example, “we should be serious, really study”), because their residues are semantically independent but of different meanings—each morpheme has independent semantics, but they are not symmetric redundancy—and recovery is more difficult than for “good-good.” Coordinative compounds (for example, “uneven”) may in theory still be penetrated by a marginal boundary under a strong complaining context (for example, “this thing is really uneven, not aligned”), but their reverse decoding relies neither on the independent semantics of the morphemes nor on symmetric redundancy; rather, it relies on phonological association and the completion of the pairing—this constitutes a third and also the most difficult type of residue. The three types, in ascending order of coverage difficulty, are: semantically independent redundant type, semantically independent but different-meaning type, and phonology-pairing-completion type. This shows that the difficulty

of coverage depends on the type, abundance, and activation conditions of the residual cues, rather than on the type of prosodic word itself. Even the most fundamental boundary internal to a prosodic word is not absolutely impassable under sufficiently extreme pragmatic conditions—its resistance is a continuous variable, not a rigid switch. But if the structure is further damaged—for example, if the word order is completely inverted as “*learn good-good we should” — then not only is the boundary internal to the prosodic word punctured, but the linear order among the constituents and the syntactic-structural relations are also completely scattered, leaving the hearer unable to find any reliable structural cues to initiate reverse decoding. At that point, even the most permissive pragmatic context cannot yield any form of acceptability. The coverage mechanism fails because “the decoding cost exceeds the communicative benefit.” This prediction makes the dual-track model not only descriptive but also falsifiable.

5. Conclusion: “a + pause” —A Window onto the Dual-Track Nature of Language Cognition

Using a set of extremely minimal Chinese minimal pairs, this paper has argued that the distributional restrictions on “a + pause” are a window for observing an effective window onto the syntax-prosody-pragmatics interface. The core findings can be summarized in two points.

First, the legality of surface pauses is not determined by a single factor, but is strictly constrained by the mapping “depth of syntactic hierarchy → level of prosodic boundary.” The three positions—after the subject, after the functional head, and within the prosodic word—constitute a clear gradient of boundary strength, and the rigid computation of the formal path has precise predictive power for this.

Second, this rigid gradient can be “covered” under specific pragmatic conditions—but such covering does not alter the underlying syntactic judgment. Rather, through the mechanism of identity reframing on the peripheral pragmatic-inference track, it realizes a systematic switch in the source of acceptability. We distinguished two subtypes—strategic covering (a production-side rhetorical violation) and tolerance-based covering (comprehension-side fault forgiveness)—in order to characterize more finely the multiple ways in which pragmatics intervenes.

The resulting “core-periphery dual-track interaction model” preserves both the structural predictive power of the formal path and the contextual sensitivity of the functional path, and reanalyzes the conflict between the two as real-time interaction between two cognitive subsystems. This interaction leaves a clearly identifiable trace at the physical point of “啊 + pause.”

The present study is still a theoretical model combined with preliminary pre-experimental validation, and three directions merit further development. First is empirical deepening: a small-scale acceptability pre-experiment with 39 par-

ticipants has been completed, preliminarily verifying the group robustness of the boundary-strength gradient and the pragmatic-covering effect; in follow-up work, the representativeness of the sample will be expanded, and prosodic parameters such as pause duration and pitch reset in different contexts will be measured through acoustic analysis, so as to further verify the residual effects of core-track computation at the level of phonetic implementation, as well as the cognitive processing costs accompanying peripheral-track covering. Second is methodological generalization: the “啊 + pause” test will be extended to other functional categories—such as the aspect markers “了,” “着,” and “过,” the structural particle “的,” and more sentence types—in order to test its reliability and validity as a general diagnostic tool for boundary strength. Third is deepening the syntax-pragmatics interface: the possible interface between the “peripheral pragmatic-inference track” and the left-peripheral structure of the CP layer/ForceP at the level of syntactic realization will be explored, so as to implement the operation of “identity reframing” as a more concrete syntax-pragmatics mapping mechanism.

If these three directions can yield results consistent with the theoretical predictions, then we have reason to believe that these four simple Chinese sentences are indeed worthy of becoming a standard case in discussions of linguistic theory.

References

- [1] Duanmu San. The Rhythm of Chinese[J]. *Contemporary Linguistics*, 2000, 2(4): 203-209.
- [2] Feng Shengli. Chinese Prosody, Morphology, and Syntax[M]. Beijing: Peking University Press, 1997.
- [3] Lü Shuxiang. Grammar Study[M]. Beijing: China Youth Publishing House, 1956.
- [4] Xiong Ziyu, Lin Maochan. The prosodic features of “啊” and its discourse-communicative functions[J]. *Contemporary Linguistics*, 2004(2): 116-127 +189.
- [5] Wang Yu, Bi Yanjuan. The modal particle “啊”: three-way classification and its forms and functions[J]. *Journal of Foreign Languages (Journal of Shanghai International Studies University)*, 2017, 40(2): 11-27.
- [6] Couper-Kuhlen E, Selting M. *Interactional Linguistics: Studying Language in Social Interaction*[M]. Cambridge: Cambridge University Press, 2018.
- [7] Friederici A. D. Towards a Neural Basis of Auditory Sentence Processing[J]. *Trends in Cognitive Sciences*, 2002, 6(2): 78-84. DOI:10.1016/S1364-6613(00)01839-8.
- [8] Jackendoff R. *The Architecture of the Language Faculty*[M]. Cambridge, MA: MIT Press, 1997.

- [9] Jackendoff R. *Foundations of Language: Brain, Meaning, Grammar, Evolution*[M]. Oxford: Oxford University Press, 2002.
- [10] Nespor M, Vogel I. *Prosodic phonology*[M]. Dordrecht: Foris Publications, 1986.
- [11] Selkirk E. O. *Phonology and Syntax: The Relation between Sound and Structure*[M]. Cambridge, MA: MIT Press, 1984.
- [12] Sperber D, Wilson D. *Relevance: Communication and Cognition*[M]. 2nd ed. Oxford: Blackwell, 1995.

Appendix: Materials for the Acceptability Pre-experiment

A Opening Instructions

Hello! This is a short survey on intuitions about spoken Chinese. There are no right or wrong answers; you only need to respond independently according to your genuine feeling about everyday speech. The entire questionnaire will take about 3-5 minutes. Thank you for your cooperation!

Response tips:

1. First read the sentence aloud softly once—or read it silently in your mind—imitating the tone of everyday speech, and then quickly judge it based on your first impression;
2. Do not repeatedly deliberate over grammatical correctness, and do not discuss your answers with others;
3. Use a 1-7 scale for scoring:
7 points: completely natural; this is exactly how one would say it in everyday speech
5 points: basically smooth; somewhat disfluent but very normal
3 points: relatively awkward; people generally would not say it this way, but the meaning is understandable
1 point: completely disfluent; it simply sounds wrong

B Practice Items (not included in the statistics)

1. This afternoon we are going to the library to read.
2. We eat the meal in the cafeteria.

C Neutral-Context Test Items (presented randomly)

Core test sentences:

1. We should study hard. (core baseline sentence)

2. We, ah, should study hard. (core—after the subject)
3. We should, ah, study hard. (core—after the functional head)
4. We should study ha, hard. (core—inside the prosodic word)

Filler sentences:

5. It will probably rain tomorrow. (Fill-in—baseline natural sentence)
6. I think this matter is not very appropriate. (Fill-in—baseline natural sentence)
7. Let's go hiking together this weekend. (Fill-in—baseline natural sentence)
8. He goes running on the sports field every morning. (Fill-in—baseline natural sentence)
9. As for him, ah, he has always had this temper. (Fill-in—sentence with “ah” after a legitimate subject)
10. As for this matter, ah, we still need to discuss it further. (Fill-in—sentence with “ah” after a legitimate subject)
11. As for the holiday, ah, I plan to go back to my hometown for a few days. (Fill-in—sentence with “ah” after a legitimate subject)
12. We must, ah, take this matter seriously. (Fill-in—awkward sentence with “ah” after another function word)
13. He can, ah, run five kilometers in one breath. (Fill-in—awkward sentence with “ah” after another function word)
14. They will, ah, complete the task on time. (Fill-in—awkward sentence with “ah” after another function word)
15. I will go to the office after eating-finished. (Fill-in—error-checking sentence)
16. This book, I finished reading yesterday. (Fill-in—error-checking sentence)

D Contextualized Test Items

Scenario 1: At a class meeting, the homeroom teacher exhorts the whole class, speaking in an earnest tone, at a relatively slow pace, and with heartfelt emphasis.

1. We must, ah, study hard and must not slack off anymore. (Control: after a function word + strategic coverage)
2. We must stu—ah—study hard! We must not keep playing around. (Core: within a prosodic word + strategic coverage)

Scenario 2: A student is somewhat nervous while speaking on stage and accidentally gets stuck mid-sentence.

1. We must, ah...study hard. (Control: after a function word + error-tolerance coverage)
2. We must stu—ah...study hard. (Core: within a prosodic word + error-tolerance coverage)

E Background Information Questions (Backup)

1. Your native-language background is: mainly Mandarin mainly dialect
(please specify: _____)
2. Have you systematically studied knowledge related to Modern Chinese/linguistics? Yes No
3. Your age group: under 18 18-25 26-40 41 and above

Corresponding author: Shi Wei (E-mail: 20031@jdzu.edu.cn)

Author Contribution Statement

Shi Wei: proposed the research idea and designed the research plan; was responsible for the literature review, construction of the theoretical framework, and methodological innovation; completed the drafting, revision, and final-version revision of the paper.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv. Machine translation. Verify with the original.