Trans-domain polygrammaticalization via recycling: The case of the multifunctional affix *ge?* in Jin Chinese*

Chenchen Song, Shangze Li & Adæmrys Chihjen Cheng

Manuscript (Sep. 2025)

Abstract

Multifunctional morphemes are common in human language and are often the synchronic reflection of diachronic linguistic changes like grammaticalization. However, there are also difficult cases of multifunctionality that cannot be easily explained by grammaticalization theory or its extensions such as polygrammaticalization and transcategoriality. In this paper, we study a "difficult" multifunctional affix of this kind: *ge?* in Jin Chinese. Its synchronic uses are not all obviously related, and its historical evolution involves back-and-forth trans-domain changes between phonology and morphosyntax. We first systematically introduce its multiple functions in a theoretically and cross-linguistically oriented manner and then develop a formal-syntactic model to derive them step by step, mainly using tools from (generalized) root syntax. After that, we apply our synchronic analysis to the diachronic aspects of the morpheme and account for its peculiar grammaticalization scenario by a domain-general notion of recycling, which has been argued to reflect a general cognitive factor in the design of human language.

Keywords: Jin Chinese, multifunctionality, grammaticalization, recycling, root syntax

1 Introduction

A basic fact about human language is that a single linguistic form may belong to several categories, thus shouldering multiple lexical or grammatical functions. For example, *that* in English can be used as a demonstrative (e.g., *that car*) or a complementizer (e.g., *he said that …*). Multifunctional items are especially common in isolating languages like Chinese. To illustrate, consider the Mandarin function word *de*, whose multiple uses are listed in (1).[1]

*This paper is the written-up and revised version of two conference presentations: "Multifunctionality across categories: The case of Gaoping Jin *ge*" (Song & Li 2024, SinFonIJA 17, Nova Gorica) and "Two levels of degree reinforcement: The case of A-*ge*-BB adjectives in Gaoping Jin" (Song, Cheng & Li 2024, SICOGG 26, Jeonju). Thanks to the audiences at both conferences for helpful feedback. The case study reported here is ultimately based on Shangze Li's MA thesis (2024, Zhejiang University), though the version presented in this paper has been substantially further developed both empirically and theoretically. Special thanks to the Cambridge recycling reading group (2024–2025) organized by Theresa Biberauer and Cora Cavirani-Pots, which has provided important inspiration for the diachronic aspect of this paper.

[1]We use the following abbreviations in our glosses: ASP = aspect, CL = classifier, COMP = complement marker, CONT = continuative, DIM = diminutive, GEN = genitive, ITE = iterative, LK = linking element, MODI = modification marker, NOM = nominative, ONOM = onomatopoeic, PFV = perfective, PRT = particle

(1)

V ‘get, obtain’	***de*** *jiang* ‘get (a) prize’	[Standard Mandarin]
pre-V modal	*bu* ***de*** *jinru* ‘not can enter; cannot enter’	
post-V modal	*ta qu-**de*** ‘he go-may; he may go’	
post-V complement marker	*ku* ***de*** *yan hong-le* ‘cry COMP eyes red-ASP’ (i.e., cry till the eyes turn red)	
post-V modal infix	*chi-**de**-wan* ‘eat-can-finish; can eat up’	

A shared property of the English and Chinese examples above is that the multiple meanings of a multifunctional word are diachronically chained by grammaticalization. The complementizer *that* historically derives from the demonstrative *that* (Roberts & Roussou 2003: 116), and Chinese *de* has the evolutionary pathway in (2).

(2) *de*: V → pre-V modal
↘ post-V modal → post-V complementizer → post-V modal infix
(L. Wang 2004, 2005)

This is more exactly a scenario of *polygrammaticalization* (Craig 1991), as there are two parallel evolutionary branches out of a single lexical source. However, since each branch still conforms to the basic definition of grammaticalization, given in (3), we may treat polygrammaticalization as a principled extension of classical grammaticalization.

(3) Grammaticalization is the linguistic change whereby a lexical item or construction takes on grammatical characteristics or through which a grammatical item becomes more grammatical. (Hopper & Traugott 2003: 2)

But there are also cases of multifunctionality that cannot be easily explained from the perspective of grammaticalization, even with its polygrammaticalization extension. The difficulty mainly stems from two tenets of grammaticalization theory: *i*) that grammaticalization is unidirectional (i.e., from less to more grammatical), and *ii*) that grammaticalization is a relation between meaningful linguistic units (i.e., grammatical meaning does not pop out of thin air).

A well-known counterexample to the unidirectional tenet is the replacement of affixal genitive inflection in Old English, illustrated in (4), by the clitic *-’s* in Modern English. Since affixes are considered more grammatical than clitics, this grammaticalization is in the opposite direction to the standardly assumed cline given in (5).

(4) *ðaes cyning-es sweoster Ecgfrið-es* [Old English]
the.GEN king-GEN sister.NOM Ecgfrið-GEN
‘the sister of Ecgfrith the king’ (Hopper & Traugott 2003: 136)

(5) content item → grammatical word → clitic → inflectional affix (ibid., p. 7)

Robert (2018) discusses nonunidirectional categorial change from a broader perspective, using the term *transcategoriality* to cover both classical grammaticalization—which she dubs “oriented transcategoriality”—and two deviating scenarios, “generic” and “functional” transcategoriality. These are respectively scenarios where an uncategorized or weakly categorized linguistic unit is directly instantiated in multiple categories (generic) and where a multifunctional grammatical morpheme or “polyfunctional” has no lexical use (functional) (pp. 127–128). The kind of “massive, polydirectional, and unmarked” (ibid.) multifunctionality commonly observed in

Chinese, among other isolating languages, is more accurately characterized by generic transcategoriality.

Counterexamples to the second tenet above is harder to imagine. Can grammatical items appear all of a sudden? Hopper and Traugott (2003: 132) remark that it is difficult to say whether grammatical items can "arise full-fledged … without a prior lexical history" due to our limited ability to look back into (pre)history. For current purposes, we can relax the standard a little and consider a grammatical item not "out of thin air" as long as it has a meaningful linguistic unit as its precursor, be that lexical or grammatical. Under this relaxation, none of the transcategorial morphemes discussed in Robert (2018) are really out of thin air.

Against the above background, in this paper we study a peculiar case of multifunctionality that violates both tenets of grammaticalization theory. On the one hand, it is not unidirectional change. On the other hand, it involves grammaticalization out of thin air. This is the case of the affix *ge?* in Jin Chinese. Jin is a non-Mandarin major variety of Chinese spoken in Shanxi Province and adjacent regions. The grammaticalization of *ge?* is partly out of thin air in that its origin is not a meaningful linguistic unit but a mere syllable. That is, its evolution is not only transcategorial but also *trans-domain*, so to speak, from phonology to morphosyntax (and it furthermore involves back-and-forth changes between the two linguistic domains). This clearly also violates unidirectionality, because the standardly assumed ultimate direction of grammaticalization is from morphosyntax to phonology, as demonstrated by the case of English *lets* in (6).

(6) a. Lets you and him fight. [Midwestern American English]

b. (let) us → (let)'s → (let)s (Hopper & Traugott 2003: 10–12)

As Hopper and Traugott show, in some varieties of English the final *s* in *lets* has lost "its status as a separate morpheme" and become a "simple phonemic constituent of a (monomorphemic) word" (ibid.). The situation of *ge?* is the exact opposite.

The multiple functions of *ge?* are listed in (7). The accompanying examples are from the Jin subvariety spoken in Gaoping (southeastern Shanxi), which is also our main source of data. Unless otherwise noted, the data in this paper are collected by ourselves.[2]

(7) a. A cross-categorial prefix for word-formation purposes (e.g., ***ge?****-duei* 'GE-pile; mound of earth', ***ge?****-chua* 'GE-poke; plug in')

b. An iterative-diminutive aktionsart verbal prefix (e.g., ***ge?****-hua* 'GE-draw; scribble')

c. An exaggerative degree infix for ABB-shaped adjectives (e.g., *ha?-**ge?**-dongdong* 'black-GE-BB; so pitch-dark!!')[3]

The morpheme *ge?* has received much previous research (see L. Zhang 2020 for an overview). However, previous studies are almost all in Chinese (except S. Li 2024, which this paper partly builds on) and hence inaccessible to the wider international audience. In addition, with one or two rare exceptions (e.g., X. Wang 2022), previous studies are all descriptive. Given this status quo, the case of *ge?* is worth further investigation both from a grammaticalization perspective and from a theoretical morphosyntactic perspective. We take on both tasks in this paper and make three contributions. First, we systematically introduce the multifunctionality of *ge?* in

[2]The second author of this paper, Shangze Li, is a native speaker of Gaoping Jin.

[3]We use "BB" as a placeholder in our gloss here because the BB part of an ABB adjective is generally iconic or ideophonic, having no concrete lexical meaning (Van Hoey 2023).

a theoretically oriented and cross-linguistically contextualized manner. Second, we develop a unified formal model to derive the three functions of *ge?* in the generative syntactic framework, mainly using tools from root syntax. Third, we further extend our synchronic analysis to the diachronic aspects of *ge?* and propose a principled formal account of its peculiar grammaticalization scenario, which we identify as a case of trans-domain polygrammaticalization via recycling.

The rest of the paper is structured as follows. In Section 2, we examine the multifunctionality of *ge?*. In Section 3, we develop our formal-theoretical model. In Section 4, we extend our synchronic analysis to the diachronic development of *ge?*. Section 5 concludes.

2 The multifunctionality of *ge?*

The affix *ge?* is a well-known element of Jin Chinese (Hou 1999: 21). In this section, we examine its three functions in detail, mainly based on data from Gaoping Jin but also citing examples from other Jin subvarieties when necessary. We begin with the word-formation prefix use of *ge?* (§2.1) and then move on to its aktionsart prefix use (§2.2) and its degree infix use (§2.3). As we will see, the latter two uses are more regular and productive, at least in some regions, whereas the first use is a fossilized phenomenon.

2.1 Word-formation prefix *ge?*

The most noticeable characteristic of the word-formation use of *ge?* is that it is cross-categorial. According to a consensus in the literature (Ma 1995; Hou 1999; Qiao 2023), *ge?* can be used in five categories: noun (N), classifier (Cl), verb (V), adjective (A), and onomatopoeic word (Onom). See (8) for some example items and see (9) for some example sentences.

(8) a. *ge?-duei*(*-r*) 'GE-pile(-N); small mound of earth' [Gaoping Jin]
ge?-dou(*-r*) 'GE-pit(-N); small pit'
ge?-ba(*-r*) 'GE-handlebar(-N); knob' (N)

b. *ge?-duo-zhe* 'GE-bulb.shaped-N; for bulb-shaped small objects'
ge?-lian-r 'GE-small.piece-N; for small pieces of land'
ge?-seng-r 'GE-star-N; for small amounts' (Cl)

c. *ge?-chua* 'GE-poke; plug in'
ge?-xian 'GE-lift.cover; remove (lid, cover, etc.)'
ge?-chuo? 'GE-wrinkle; wrinkle, crumple' (V)

d. *ge?-liou* 'GE-bent; (slightly) bent'
ge?-zha 'GE-dreg; dirty'
ge?-chuo? 'GE-wrinkle; wrinkled' (A)[4]

e. *ge?-dong* 'thumping sound'
ge?-cha 'cracking sound'
ge?-zhi 'creaking sound' (Onom)

[4]Note that *ge?-chuo?* is both a verb and an adjective. Such categorial flexibility is common in Chinese.

(9) a. *Va kanzhae yeu ji ge* ***ge?-duei-r****, sousin sengr ba!* [Gaoping Jin]
I see have several CL GE-pile-N careful a bit SFP
'I saw several small mounds of earth (over there), so be careful!' (N)

b. *Ge va na yi?* ***ge?-duo-zhe*** *shuan!*
give me bring one GE-bulb.shaped-N garlic
'Get me a small head of garlic!' (Cl)

c. *Va diou chongdianqi cheinei yi?zhe* ***ge?-chua*** *bu jinge lêi?!*
I this charger why always GE-poke not enter SFP
'Why can't I successfully plug in my charger (despite many attempts)?!' (V)

d. *Na shang nian genr* ***ge?-liou*** *zhizhe!*
pick up that CL GE-bent twig
'Pick up that (slightly) bent twig!' (A)

e. *Ni zhêi nar* ***ge?-dong ge?-dong*** *nong shein lêi?*
you be at there GE-ONOM GE-ONOM do what SFP
'What are you doing there making that thumping noise?' (Onom)

We can make several observations regarding the word-formation prefix *ge?*. First, this prefix is clearly derivational, not inflectional, as a *ge?*-prefixed word and its base may denote different lexical concepts. In some cases, such as the adjective *ge?-liou* (8d), *ge?* is not really a result of prefixation but one of syllable-splitting, where an originally monosyllabic word (here *gjiew* 'bend, curl' from Middle Chinese) is split into two syllables. Despite their diachronic differences, however, synchronically all word-initial *ge?*s are perceived by ordinary speakers as prefixes.[5] As we will see in Section 4, split-syllable words are the historical origin of *ge?*.

Second, the nouns and classifiers in (8a/b) all carry an extra suffix *-r*/*-zhe*. These are the Gaoping Jin versions of the general nominal markers in Modern Chinese (corresponding to Mandarin *-er*/*-zi*). In Gaoping Jin, such a suffix frequently appears on nouns and classifiers, more obligatorily on the latter. Such nominal suffixes can appear on classifiers because most classifiers are derived from nouns (L. Wang 2001; H. Fan 2007).[6] Historically, these suffixes used to have diminutive meaning, but their diminutivity has been much eroded now due to further lexicalization (X. Wang 2022: 241). We simply gloss them as N for expository convenience. That these suffixes are a separate phenomenon, independent of *ge?*, is evidenced by the fact that they also occur on words without *ge?*, as in *men-r* 'door-N', *ben-r* 'volume-N; classifier for books', *shuo?-zhe* 'bundle-N; classifier for flowers', and so on.

Third, *ge?* is often said to be a diminutive affix itself (Ma 1995; Hou 1999; H. Fan 2007), which is also a general intuition of native speakers. Indeed, all the nouns and classifiers in (8) are diminutive, and some of the verbs and adjectives are somewhat diminutive too. Thus, *ge?-liou*, *ge?-chua*, and *ge?-chuo?* in (8c/d) can be considered diminutive in the sense that they denote intrinsically "small" concepts (see also Xing 2002: 268–269), though *ge?-xian* and *ge?-zha* cannot be considered in this way. In fact, we could say the same thing about the nouns and classifiers in (8), where the diminutive reading is also often an intrinsic property of the lexeme concept rather than a contribution of *ge?*. Even some *ge?*-prefixed onomatopoeic words, such as *ge?-zhi* in (8e), may be considered diminutive in this way—by mimicking subtle sounds.

[5] Native speakers are not quite aware which instances of *ge?* are real prefixes and which are not, and neither are many linguists, as pointed out by P. Bai (1988). This probably means that all items with the word-formation *ge?* are synchronically represented in the same way in speakers' minds.

[6] In fact, the category of classifiers itself developed from nouns in history (L. Wang 2004, 2005).

Hence, the real pattern here seems to be that *ge?* tends to be selected by roots denoting "small" concepts, not that it marks diminutivity itself (see also H. Fan 2007). Indeed, there is no good explanation as to why the items in (8) require *ge?* but those in (10) repel it, other than that this is an idiosyncratic requirement of certain roots.

(10) a. (**ge-*)*men-r* 'GE-door-N', (**ge-*)*hua-r* 'GE-flower-N' (N) [Gaoping Jin]

b. (**ge-*)*jian-r* 'GE-piece-N; for clothes', (**ge-*)*ben-r* 'GE-volume-N; for books' (Cl)

c. (**ge-*)*zheu* 'GE-walk', (**ge-*)*kan* 'GE-look' (V)

d. (**ge-*)*ha?* 'GE-black', (**ge-*)*tian* 'GE-sweet' (A)

e. (**ge-*)*seu* 'GE-whoosh', *(*ge-)pu* 'GE-puff' (Onom)

There is cross-regional variation as to which roots require *ge?*, and in some regions *ge?* seems to have developed authentic diminutive-marking function in adjectives, as in (11).[7]

(11) *ge?-hong* 'GE-red; reddish', *ge?-ruan* 'GE-soft; softish' [Wenshui Jin] (Ma 1995: 41)

In sum, the word-formation prefix *ge?* in Jin Chinese is idiosyncratically selected by a subset of roots to form new lexemes and is not restricted by category. It tends to be selected by roots denoting "small" concepts and may sometimes contribute diminutive meaning itself.

A direct consequence of the lexeme status of *ge?*-prefixed words is that they can be input to further morphological processes. For instance, *ge?*-prefixed words of various categories can be partially reduplicated, as in (12). The semantic effect of such reduplication is category-dependent.

(12) a. *ge?-tai* 'stairs, steps' (N) → *ge?-taitai* 'small stairs (diminutive)' [Xinfu Jin]
ge?-ni 'oily' (A) → *ge?-nini* 'very oily (augmentative)'

b. *ge?-kan* 'take a look' (V) → *ge?-kankan* 'take a tiny look' [Fenyang Jin]
ge?-nuo 'move a bit' (V) → *ge?-nuonuo* 'move a tiny bit' (Qiao 2023: 3)

In particular, reduplicated *ge?*-verbs in Fenyang Jin have a doubly diminutive effect. This is presumably because the *ge?* in such verbs is the aktionsart *ge?* to be discussed in Section 2.2.

Similarly, *ge?*-prefixed words may also be input to compounding. For example, many *ge?*-prefixed nouns can be compounded with another noun, yielding the N_1-*ge?*-N_2 construction, as in (13).

(13) *tu-ge?-duei(-r)* 'earth-GE-pile(-N); small earth mound' [Gaoping Jin]
shan-ge?-va(-r) 'mountain-GE-den(-N); out-of-the-way mountain area'
rou-ge?-dun(-r) 'meat-GE-pillar(-N); someone chubby in a cute way'

2.2 Aktionsart prefix *ge?*

The *ge?*-prefixed verbs we have seen so far all have idiosyncratic meanings unpredictable from their bases. But there are also *ge?*-prefixed verbs with more regular and predictable meanings, as in (14).

[7] All data cited from the literature have been adapted to match the style of this paper.

(14) *ge?-hua* 'GE-draw; scribble' [Gaoping Jin]
ge?-shan 'GE-fan; keep fanning (to make breeze)'
ge?-fan 'GE-search; dig around'

The first thing we can notice about these *ge?*-prefixed verbs is that their bases are all standalone verbs with clearly defined meanings. This situation is very different from that of the *ge?*-prefixed verbs from Section 2.1, which often involve bases that cannot be used alone or have no clear meaning (e.g., **chuo?* in *ge?-chuo?* 'wrinkle, crumple'). Semantically, the *ge?* in (14) consistently means 'do something with reduced force but increased frequency' or 'repeatedly but slightly do something within a short time'. See (15) for some example sentences.

(15) a. *Buyou zhêi tsang shang* ***ge?-hua****!* [Gaoping Jin]
don't be at wall on GE-draw
'Don't scribble (i.e., casually draw) on the wall!'

b. *Ge va* ***ge?-shan****-zhu dianr feng!*
give me GE-fan-dwell.CONT a bit wind
'Keep fanning me (to help me cool down)!'

c. *Jizhu buyou zhou ta qi va nar* ***ge?-fan****-ge a!*
remember don't let him go I there GE-search-thither SFP
'Remember not to let him go digging around (i.e., casually searching) in my room!'

Note that *ge?*-prefixed verbs can normally take verbal complements, such as the continuative *zhu* 'lit. dwell' and the perspectival *ge* 'lit. go', which are a common means to form complex verbal predicates in Chinese.

The above-described regular meaning of *ge?* matches the meaning of the iterative-diminutive aktionsart in Kiefer (2010). According to Kiefer, aktionsart is "the modification of verb meaning by morphological means" through "add[ing] one or two semantic features to the meaning of the base verb" (p. 145). Iterative and diminutive are two aktionsarten that are often tied to each other, as illustrated by the cross-linguistic examples in (16).

(16) a. *po-čit-yvat'* 'DIM-read-ITE; read a little from time to time' [Russian]

b. *néz-e-get* 'look-DIM; look a little bit' [Hungarian]
nyit-o-gat 'open-ITE; open repeatedly' (Kiefer 2010: 149, 157, 161)

In (16a), the Russian diminutive and iterative affixes are applied to the same root *čit* 'read'. In (16b), Hungarian *-get*/*-gat* can denote either the diminutive or the iterative aktionsart.[8] Similarly, in Jin Chinese the diminutive and iterative aktionsarten are both realized by *ge?*.

Unlike the word-formation *ge?*, which is unproductive in the entire Jin-speaking area, the aktionsart *ge?* is very productive in some regions, including Shenmu (Xing 2002) and Pingyao (Hou 1999; X. Wang 2022).[9] However, it is not that productive in Gaoping. According to T. Zhang (2021), in Gaoping Jin the aktionsart *ge?* is mostly limited to high-frequency activity verbs denoting concrete actions. T. Zhang gives the further examples in (17a) and remarks that *ge?* cannot be used with the verbs in (17b).

[8]The vowel alternation here is due to vowel harmony, and *e*/*o* in (16b) are epenthetic vowels. According to Kiefer (1996: 180–181), this suffix can express iterativity, diminished intensity, lack of precision, etc.

[9]According to Ma (1995: 40), the aktionsart *ge?* has the highest level of productivity in Central Jin.

(17) a. *ge?-deng* 'GE-wait; wait a bit', *ge?-mi* 'GE-nap; nap a bit' [Gaoping Jin]
ge?-qi 'GE-ride; ride a bit', *ge?-shuan* 'GE-rinse; rinse a bit'

b. **ge?-duo?* 'GE-read', **ge?-shuei* 'GE-sleep'
**ge?-lêi* 'GE-come', **ge?-êi* 'GE-love' (T. Zhang 2021: 5–7)

Overall, it seems that in Gaoping Jin the set of verbs that can take the aktionsart *ge?* is idiosyncratically defined. That is unsurprising since aktionsart-marking, albeit compositional, is still a process of derivational morphology.

There are also ambiguous *ge?*-prefixed verbs in some Jin subvarieties, depending on whether the *ge?* in question is in the word-formation or the aktionsart use, as in (18).

(18) a. *ge?-me?* 'GE-touch' [Panshang Jin]
(i) walk slowly (back and forth); (ii) casually touch

b. *ge?-jiao* 'GE-stir'
(i) disturb; (ii) casually or slightly stir (X. Wang 2022: 71)

Between the two meanings of each verb in (18), the first is idiosyncratic (i.e., a new lexical meaning), while the second is compositional.

2.3 Degree infix *ge?*

Now we turn to the third use of *ge?*—as an exaggerative degree infix. This use is limited to state-denoting adjectives in the ABB form, where A is the base (mostly adjectival[10]), encoding a basic state, and BB is an idiosyncratically selected depictive element in a reduplicated form, which vividly augments the degree of that state. See (19) for some examples and note how *ge?* further augments the degree of ABB in an expressive and emotional way.[11]

(19) ABB → A-*ge?*-BB [Gaoping Jin]
ha?-dongdong 'black-BB; pitch-dark' → *ha?-ge?-dongdong* 'so pitch-dark!!'
nan-yiengyieng 'blue-BB; brightly blue' → *nan-ge?-yiengyieng* 'so brightly blue!!'
tian-zi?zi? 'sweet-BB; pleasantly sweet' → *tian-ge?-zi?zi?* 'so pleasantly sweet!!'

ABB adjectives widely exist in Sinitic languages (e.g., Standard Mandarin *hei-qiqi* 'black-BB; pitch-dark', Shanghai Wu *ku-jiji* 'bitter-BB; very bitter'). Admittedly, an ABB adjective is already expressive and emotional—and evaluative too according to Dong (2021)[12]—but adding *ge?* makes it even more so. In other words, there are two levels of morphological degree reinforcement in Jin Chinese: a state-denoting A can first be expanded to ABB and then be further expanded to A-*ge?*-BB. See (20)–(21) for some example sentences.

[10] The real condition here, in Sinitic languages in general, is that A denote a state (C.-S. L. Liu 2013). Nouns and verbs may occasionally fill the slot too as long as they are temporarily used to denote states, as in Mandarin *lei-wangwang* 'tear-BB; teary-eyed' and *xiao-haha* 'laugh-BB; laughing heartily', where the noun *lei* and the verb *xiao* temporarily denote a tearing-up state and a laughing state.

[11] Recall from note 3 that we gloss BBs with the placeholder "BB" due to their ideophonic nature (Van Hoey 2023). We let their vivid effects be directly reflected in the free English translations. This can also help us avoid wrong lexical meaning assignments induced by misleading Chinese characters, which are often only used to record the sounds of BBs. See Shao (1990), Dong (2021), and Van Hoey (2023) for discussions.

[12] Native speakers can feel what the effect of each BB is, but the feeling is ineffable. Dong (2021: 285) attributes this ineffability or untranslatability to the evaluative nature of BBs. Evaluatives are a species of expressives (Fortin & Rainer 2022), which generally manifest "descriptive ineffability" according to Potts (2007: 177).

(20) a. *Vêiteu* ***ha?*** *la.* [Gaoping Jin]
outside black SFP
'It is dark outside.'

b. *Vêiteu* ***ha?-dongdong*** *zhe!*
outside black-BB PRT
'It is pitch-dark outside!'

c. *Vêiteu* ***ha?-ge?-dongdong*** *zhe!!*
outside black-GE-BB PRT
'It is so pitch-dark outside!!'

(21) a. *Tian si* ***nan*** *zhe.*
sky is blue PRT
'The sky is blue.'

b. *Diou tian* ***nan-yiengyieng*** *zhe!*
the sky blue-BB PRT
'The sky is brightly blue!'

c. *Diou tian* ***nan-ge?-yiengyieng*** *zhe!!*
the sky blue-GE-BB PRT
'The sky is so brightly blue!!'

The two levels of degree reinforcement are also acknowledged in the literature. L. Wang (2017: 339) points out that the degree of darkness is increasingly high in the three boldfaced words from (20). Tian (2015: 21) and Qiao (2023: 53) make similar remarks based on other Jin subvarieties. This infixal use of *ge?* is productive in Gaoping Jin (in our observation) and Pingding Jin (Huo 2015: 39) but nonproductive in, for instance, Shenmu Jin (Xing 2002: 528).

Above we have only seen A-*ge?*-BB items used predicatively, but they can be used modificationally too (22). However, they cannot normally co-occur with degree adverbs (23).

(22) a. *Ta* ***ze?-ge?-huahua*** *zhe pou guolêi la!!* [Gaoping Jin]
she hurry-GE-BB MODI run hither SFP
'She ran here in such a hurry!!' (adverbial)

b. *Zheu ta niong* ***nan-ge?-due?due?*** *niong yer, neng gancheng ge shein?!*
just he that lazy-GE-BB that look can accomplish CL what
'With his lazy-as-hell look, what can he ever accomplish?!' (attributive)

(23) a. **Vêiteu hein* ***ha?-ge?-dongdong*** *zhe!!*
outside very black-GE-BB PRT
Intended: 'It is very so pitch-dark outside!!'

b. **Vêiteu tian xian* ***ha?-ge?-dongdong*** *zhe la!!*
outside sky a little black-GE-BB PRT SFP
Intended: 'It is a little so pitch-dark outside now!!'

Finally, note that though this use of *ge?* is very productive in Gaoping Jin, it still cannot be applied to just any ABB item. Since the infixation of *ge?* relies on the morphemic boundary between A and BB, fully onomatopoeic (i.e., monomorphemic) ABB items cannot take *ge?*, as in (24a). And since the A-*ge?*-BB construction is highly colloquial, it normally cannot be used on high-register ABBs either (usually Mandarin loanwords), as in (24b).

(24) a. *hualala* 'heavy rain sound' → **hua-geʔ-lala* [Gaoping Jin]
honglonglong 'heavy thunder sound' → **hong-geʔ-longlong*

b. *eʔ-heinhein* 'evil-BB; ferocious and relentless' → **eʔ-geʔ-heinhein*
seng-chongchong 'joy-BB; excited and joyful' → ?*seng-geʔ-chongchong*

But overall, at least in Gaoping Jin, this infixal use of *geʔ* is the closest to a regular inflectional affix among all its three uses we have examined.

2.4 Interim summary

To recapitulate, the three functions of Jin Chinese *geʔ* are: *i*) a cross-categorial word-formation prefix; *ii*) an aktionsart verbal prefix; and *iii*) a degree infix for ABB adjectives. There is cross-regional variation in all three functions, but overall, (i) is nonproductive in the entire Jin-speaking area, while (ii)/(iii) are productive in some regions.

An important question to ask at this point is how we can be sure that we are dealing with the same *geʔ* in all three functions. That is, how can we be sure that what we see is a case of multifunctionality rather than accidental homophony? One piece of evidence supporting our assumption that this is multifunctionality is a special situation of allomorphy. As pointed out by Xing (2002) and H. Fan (2007), among others, *geʔ* has two common allomorphs *beʔ* and *hueʔ*, which occur in both the prefixal and the infixal position, as in (25).

(25) a. Prefix: *geʔ*/*hueʔ-yo* 'slightly shake', *geʔ*/*hueʔ*/*beʔ-jio* 'slightly stir' [Shenmu Jin]

b. Infix: *tian-geʔ*/*hueʔ*/*beʔ-jiji* 'sweet-BB; so sweet!!' [Dingxiang Jin]
(Xing 2002: 276; H. Fan 2007: 43))

Among the three allomorphs, *geʔ* is the most productive, and at least in the degree infix use, the three allomorphs have developed different evaluative effects: A-*geʔ*-BB items are mostly positive—or at least nonnegative—while A-*hueʔ*/*beʔ*-BB items are mostly negative (Xing 2002; Yang 2008; Yin 2016). Thus, while all three allomorphs can be used in (25b), they reflect different speaker attitudes (i.e., whether the sweetness is perceived positively or negatively). See (26) for some further examples from Gaoping Jin.

(26) *sang-geʔ*/**hueʔ*/**beʔ-peinpein* 'fragrant-GE-BB; so pleasantly fragrant!!' [Gaoping Jin]
haʔ-geʔ/**hueʔ*/**beʔ-dongdong* 'black-GE-BB; so pitch-dark!!'
huei-beʔ/?*geʔ-chuoʔchuoʔ* 'gray-BE-BB; so unpleasantly gray and dark!!'
sêʔ-hueʔ/**geʔ-lala* 'blood-HUE-BB; so unpleasantly bloody and gory!!'

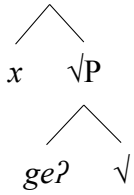
Note that the evaluative effects are already present in the BB part of these items, so what we see is probably a selectional effect—*geʔ* is the default, while *hueʔ*/*beʔ* are idiosyncratically selected by certain BBs (mostly negative). In any case, the shared allomorphy across the three functions of *geʔ* is evidence that we are faced with a scenario of multifunctionality, not accidental homophony.

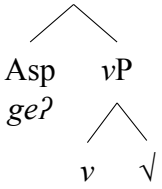
3 A formal model

In this section, we present our formal-theoretical analysis of Jin Chinese *geʔ*, modeling its three functions within the generative syntactic framework. We proceed in the same order as in the previous section, deriving the word-formation *geʔ* (§3.1), and aktionsart *geʔ* (§3.2), and the degree *geʔ* (§3.3) in turn.

3.1 Deriving the word-formation *ge?*

The only existing formal-theoretical work on *ge?* to our knowledge is X. Wang (2022), where the word-formation *ge?* is analyzed as a pre-categorization root-augmenting element in a model within Distributed Morphology (DM) (Halle & Marantz 1993, et seq.), as in (27a), where *x* is a categorizer and √ is an acategorial root. The aktionsart *ge?*, on the other hand, is analyzed as an aspectual head above an already categorized verb, as in (27b).

(27) a. 

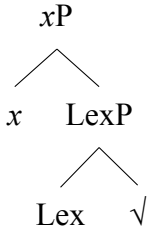
b. 

(adapted from X. Wang 2022: 258)

We first examine the derivation of the word-formation *ge?* in (27a) and will return to that of the aktionsart *ge?* in (27b) in Section 3.2. X. Wang does not discuss the degree infix *ge?*.

The two key assumptions in (27a) are that the word-formation *ge?* is categorially underspecified and that the cross-categorial nature of *ge?*-words is due to the variable categorizer *x*. While sympathetic to such an underspecification approach, we think that its implementation could be refined, for the following reasons. First, it is unclear why, between the root and *ge?*, it is the root that projects. The two nodes are in a symmetric relation in (27a), so justification is needed whichever of them is identified as the head. Second, the place of *ge?* in the syntactic category system is unclear. Being categorially underspecified in a DM model means that it is like a root, but X. Wang does not further address this issue. Third, it is unclear how the prefixal status of *ge?* can be derived from the symmetric structure in (27a). X. Wang does not address this either, though it does need an explanation considering similarly positioned root-augmenting elements in other languages tend to be suffixes (see below). Fourth, since there are only three categorizers in DM (*n*, *v*, *a*), this implementation leaves *ge?*-prefixed classifiers and onomatopoeic words unexplained. In what follows, we refine X. Wang's approach by overcoming these problems.

Root-augmenting elements are not unique to Chinese. De Belder et al. (2014) study similar elements in other languages and propose a similarly pre-categorization position for them, dubbed Lex, as in (28).

(28) 

(adapted from De Belder et al. 2014: 151)

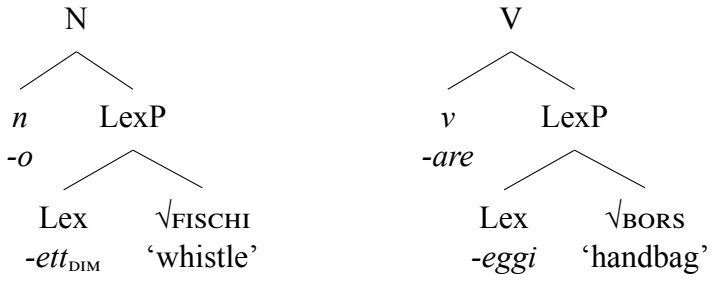
The key difference between (28) and (27a) lies in the categorial status of the root-augmenting element. While X. Wang (2022) does not assign it any category, De Belder et al. (2014) apparently assign it a category Lex and let it project—though this practice contradicts their description that Lex "has no category of its own" (p. 155) and is "similar to a root" (p. 162). As such, the categorial status of the root-augmenting element is not really clear in De Belder et al.'s model either.

As for the specific function of Lex, while it mostly contributes diminutive meaning in De Belder et al.'s data, they allow it more flexibility in principle, stating that "it is not exclusively for diminutives … [but] is more generally a position for root augmentation" (p. 163). See the Italian examples in (29).

(29) *telefon-in-o* 'telephone-LEX$_{DIM}$-N; cell phone' [Italian]
fischi-ett-o 'whistle-LEX$_{DIM}$-N; whistle (the object)'
fischi-ett-are 'whistle-LEX$_{DIM}$-V; to whistle repeatedly'
bors-eggi-o 'handbag-LEX-N; the act of mugging'
bors-eggi-are 'handbag-LEX-V; to mug' (De Belder et al. 2014: 149–163)

Like Jin Chinese *ge?*, these Italian root-augmenting elements create new lexemes and are cross-categorial,[13] though their cross-categoriality is more limited than that of *ge?* (only in N and V). Semantically, *-in* and *-ett* are diminutive, while *-eggi* adds no extra meaning. These two scenarios are illustrated by the tree diagrams in (30).

(30) a. *fischi-ett-o* 'whistle' b. *bors-eggi-are* 'to mug'



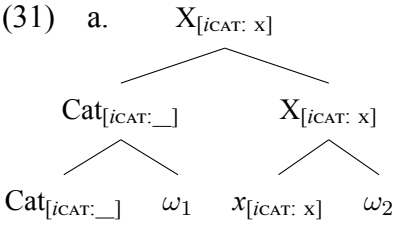
(based on De Belder et al. 2014: 159)

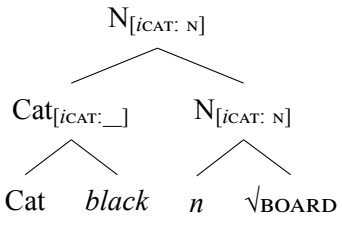
Overall, De Belder et al.'s (2014) model provides us with two useful ideas: *i*) that the root-augmenting element may be assigned some quasi-functional category while maintaining its rootlike status, and *ii*) that it may bear extra features. Meanwhile, its problem is still the unclear place of the root-augmenting element in the syntactic category system, as mentioned above.

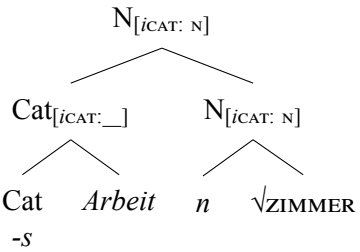
To overcome the categorial status problem, we use the "defective category" tool from Song (2019, 2025), which additionally gives us a way to derive the prefixal/suffixal status of root-augmenting elements. In his attempt to axiomatize the generative syntactic category system, Song (2019: 43ff.) proposes that in principle human language has free access to a defective categorizer, dubbed Cat, which bears an interpretable but unvalued categorial feature [*i*CAT:_]. The interpretation of Cat—as well as anything introduced by it—depends on agreement with a valued categorial feature, and upon agreement Cat gains the status of an adjunct by labeling. Song mainly uses Cat to derive modifier-head compounds: (31a) is the basic schema, and (31b/c) show the derivations of two compound words, English *blackboard* and German *Arbeit-s-zimmer* 'work-LK-room'.[14]

[13] Although De Belder et al. do not show it, the suffix *-in* is cross-categorial too; e.g., *telefon-in-are* 'to call with a cell phone' (Gaeta 2013).

[14] Song labels categorization structures by traditional "big Xs" rather than *x*Ps, highlighting the intuition that a categorization structure corresponds to a traditional content word.

(31) a. 

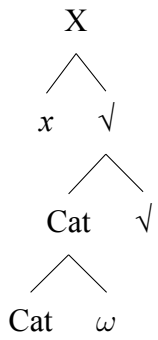
b. 

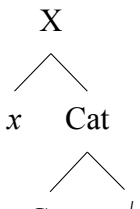
c. 

(adapted from Song 2025: 141–142)

In (31a), ω_1 and ω_2 are two syntactic objects to be (re)categorized. For modifier-head compounds, ω_1 is typically a recycled element, which can but does not have to be an existing word (e.g., *cran-* in *cran-berry*), and ω_2 is typically a root.[15] The outcome of the Cat-ω_1 merger (Cat_{ω_1} for short) further merges with X, and the outcome, having two nonminimal categories, is labeled by their shared feature [*i*CAT: x] upon Agree per the labeling algorithm (Chomsky 2013, 2015). This creates a two-segment X, making Cat_{ω_1} an adjunct (May 1985; Chomsky 1986). Then, with Cat_{ω_1} being an adjunct, the Cat_{ω_1}-X merger is linearized as $\text{Cat}_{\omega_1} \prec \text{X}$ (i.e., ω_1 becomes a "prefix") following Kayne (1994). Cat can either remain silent or be spelled out in language-specific ways, just as little *x* categorizers can. For instance, in (31c), Cat is realized as a linking element *-s*.

We can apply Cat to root-augmenting elements too, with two minor adjustments. First, instead of applying it at the X-level, we apply it at the root level. Second, instead of always letting it introduce a recycled element, we assume two possibilities: either having it introduce a recycled element or not. In the latter scenario, Cat becomes a semicategorizer for the root of X prior to the final categorization step—"semi" in the sense that Cat merely provides a categorial feature attribute but no value. We illustrate the two scenarios in (32).

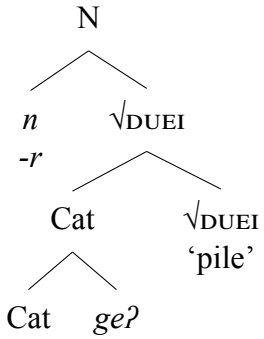
(32) a. 

b. 

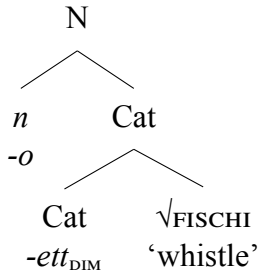
In (32a), ω can be any linguistic material as mentioned above, be it a prederived syntactic object or a primitive element (e.g., a root). Cat_{ω} merges with the root √, and the outcome is algorithmically labeled by √ since that is the minimal category between the two nodes. Of course, this labeling step is vacuous due to the acategoriality of the root, but the two √-nodes are still trivially featurally identical (by having no features); therefore, Cat_{ω} is still mechanically identified as an adjunct (i.e., a root adjunct), whence its prefixal status. By comparison, in (32b), Cat does not introduce any extra material, but since it merges with a root, we get a root categorization situation. Thus, Cat labels the Cat-√ merger and behaves morphophonologically like a normal categorizer (e.g., as a suffix in Italian). To illustrate, we derive Jin Chinese *geʔ-duei-r* 'small mound of earth' and Italian *fischi-ett-o* 'whistle (the object)' using Cat in (33).

[15] Though again not necessarily so; e.g., *ice cream* in *vanilla ice cream*, *-mas* in *Newtonmas*.

(33) a. *ge?-duei-r*



b. *fischi-ett-o*



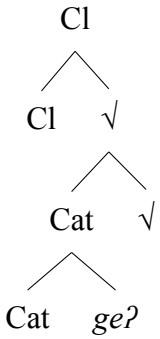
In (33a), *ge?* enters syntax by merging with Cat as a primitive element. We identify this *ge?* as a root in the spirit of X. Wang (2022) and De Belder et al. (2014) and will further motivate this theoretical choice from a diachronic perspective in Section 4.3. Next, Cat$_{ge?}$ merges with and thereby augments the root √DUEI, and the "compound root" as a whole gets nominalized by *n*, taking on the suffix *-r* as a result and yielding the linear order *ge?*≺*duei*≺*r*. On the other hand, in (33b) the simple root √FISCHI first merges with Cat, taking on the root-augmenting material *-ett*$_{DIM}$, and then the augmented root is fully categorized by *n*, taking on the nominal ending *-o*,[16] thus yielding the linear order *fischi*≺*ett*≺*o*. Our analysis of *fischi-ett-o* minimally differs from De Belder et al.'s (2014) analysis in our use of Cat instead of Lex. The Cat-based analysis has two advantages. Theoretically, Cat has a clearly defined place in the syntactic category system. Empirically, Cat provides a unified account of cross-linguistic root-augmenting elements, including their prefixal/suffixal status.

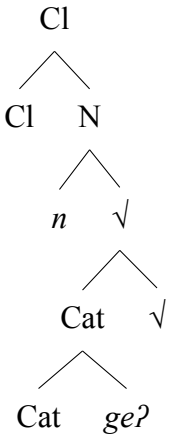
So far, we have overcome three of the four problems in X. Wang's (2022) model: First, between *ge?* and the (main) root, the latter projects because it is the minimal-category node in our Cat-based compound root configuration. Second, *ge?* is unambiguously identified as a root in our model—and so is indeed categorially underspecified—but we have externally equipped it with a formal category Cat in the meantime and thereby licensed its legal presence in syntax. Moreover, this extrinsic category does not affect the "auxiliary" status of *ge?* or accidentally let it project. Third, the prefixal status of *ge?* follows from the adjunct status of Cat$_{ge?}$ in our proposed structure for *ge?*-words.

Now, only one problem remains: none of the approaches we have seen can derive *ge?*-prefixed classifiers or onomatopoeic words. This is because categorizers are strictly defined as little *x* heads in DM. However, as Song (2019: 102) points out, that is a stipulation rather than a logical sequitur. If the formal purpose of root categorization is to equip an acategorial root with a category and thereby license its legal presence in narrow syntax, then technically any functional category can do the job, not just the little *x*s. The latter are only special in that they create traditional content words; however, those are not the only types of words in human language with idiosyncratic content, and extending root theory to the functional domain can help us model the widespread phenomenon of semilexicality (see also Acedo-Matellán & Real-Puigdollers 2019; Cavirani-Pots 2020). Adopting this generalized view on categorization—which Song (2019) calls "generalized root syntax"—we can derive *ge?*-prefixed classifiers using the same Cat-based method from above. That said, the derivation of classifiers is slightly more complicated. Recall from Section 2.1 that *ge?*-prefixed classifiers in Jin Chinese are mostly derived from nouns, which is evidenced by the fact that they mostly carry the nominal endings *-r*/*-zhe*. Thus, in addition to the direct application of the Cat-schema (34a), we also need a recategorization-based

[16] The assumption here is that *-ett*$_{DIM}$ spells out Cat just as *-o* spells out *n*.

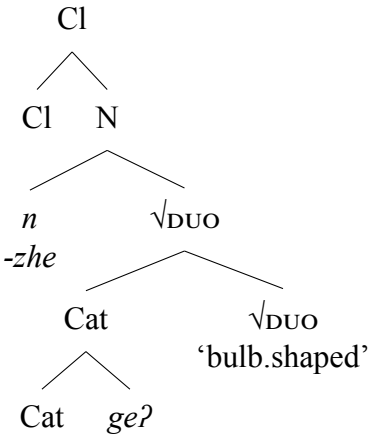
structure (34b).[17]

(34) a. 

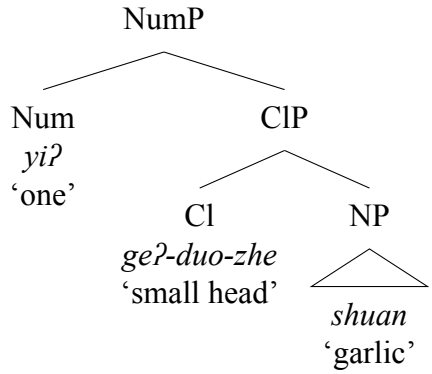
b. 

Note that in these structures, Cl is used as a categorizer rather than a functional head in the extended projection of a noun phrase (compare with (35b) below). We assume that the derivation of a *ge?*-prefixed classifier takes place in a separate workspace, which is then merged into the main workspace as a Cl head, as in (35b).

(35) a. Workspace 1
ge?-duo-zhe
'GE-bulb.shaped-N'



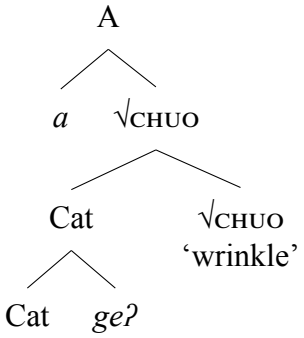
b. Main workspace
yi? ge?-duo-zhe shuan
'one small head of garlic'



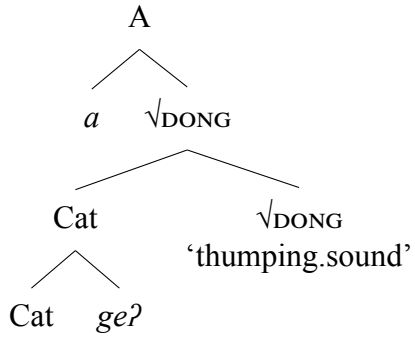
Finally, let us turn to *ge?*-prefixed onomatopoeic words. We remain agnostic as to whether or not there exists a dedicated functional category Onom but observe that onomatopoeic words are only onomatopoeic in terms of their etymology. Syntactically, they are normally used in more familiar categories (mostly A) as we have seen in Section 2.1. Hence, we derive them in the same way as *ge?*-prefixed adjectives, as in (36) (Onom-A = onomatopoeic adjective).

[17]We still keep (34a) as a theoretical possibility since dedicated *ge?*-prefixed classifiers may also exist in principle.

(36) a. *ge?-chuo?*
‘wrinkled’ (A)

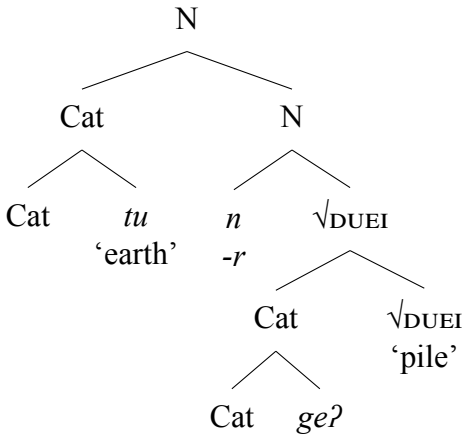


b. *ge?-dong*
‘thumping’ (Onom-A)



Before we end this section, recall from Section 2.1 that *ge?*-prefixed words can be input to further word-formation operations. This possibility is also supported by our model. Take the compounding pattern N_1-*ge?*-N_2. We derive this pattern by applying Cat at two levels, as in (37), where *tu* ‘earth’ is a noun recycled by Cat as a modifier.

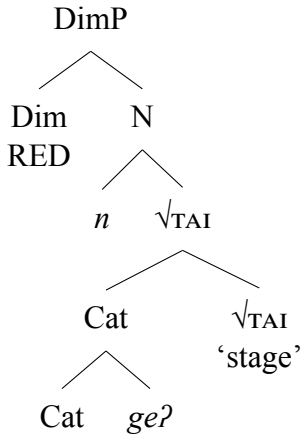
(37) *tu-ge?-duei-r* ‘small earth mound’



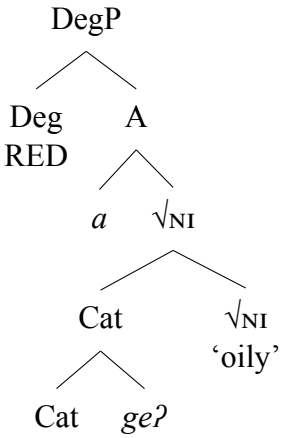
Also consider the reduplication pattern *ge?*-XX. According to X. Wang (2022), this pattern is derived by copying the last syllable of a *ge?*-word. Recall from Section 2.1 that the exact semantic effect of such reduplication varies from category to category. We assume that this is because the reduplication strategy is used to realize different functional heads, as in (38).[18] We will return to the slightly different situation in (38c) in the next section.

[18]The label Dim in the nominal case is inherited from X. Wang (2022), but one could also use De Belder et al.’s (2014) label Size.

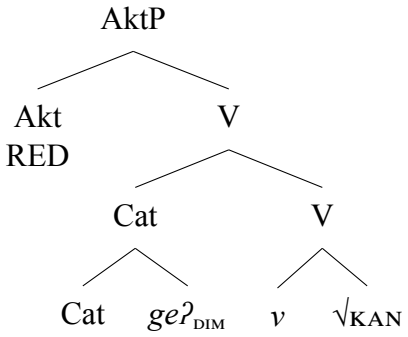
(38) a. *ge?-taitai*
‘small steps’



b. *ge?-nini*
‘very oily’



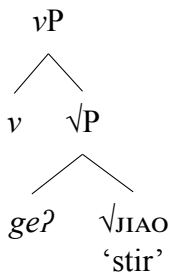
c. *ge?-kankan*
‘take a tiny look’



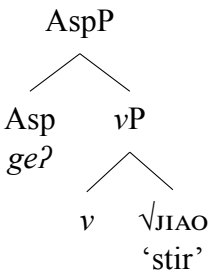
3.2 Deriving the aktionsart *ge?*

With the schema for the word-formation *ge?* in place, we can build on it to analyze the two category-specific uses of *ge?*. We begin with the aktionsart verbal prefix *ge?* in this section. As mentioned in Section 3.1, X. Wang (2022) analyzes this *ge?* as an aspectual head above a categorized verb (27b), thus derivationally distinguishing it from the word-formation *ge?*. Recall from Section 2.2 that in some Jin subvarieties there are ambiguous *ge?*-prefixed verbs. On X. Wang's analysis, this is reduced to a difference in the position of *ge?*, as in (39).

(39) a. *ge?-jiao* ‘disturb’



b. *ge?-jiao* ‘casually stir’



(adapted from X. Wang 2022: 258)

We have already improved the derivation in (39a) in the previous section. Now we examine (39b). X. Wang is right is placing the aktionsart *ge?* above the category-defining head, but the implementation in (39b) faces two problems. First, aspectual heads in Modern Chinese tend to be realized as suffixes instead of prefixes, as illustrated in (40).

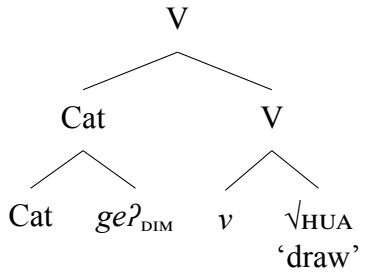
(40) a. *Baba mai-**le** xigua.* [Standard Mandarin]
dad buy-PFV watermelon
‘Dad bought watermelon.’ (grammatical aspect)

b. *Ge va ge?-shan-**zhu** dianr feng!* [Gaoping Jin]
give me GE-fan-dwell.CONT a bit wind
‘Keep fanning me (to help me cool down)!’ (lexical aspect, repeated from (15b))

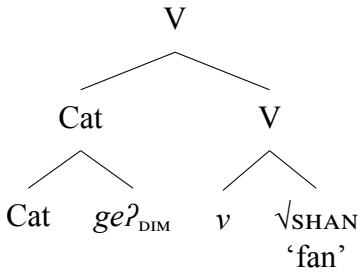
Given this general situation, analyzing the aktionsart prefix *ge?* as an aspectual head would require extra justification. Second, as we have seen in Section 2.1 and (38c), in some Jin subvarieties the aktionsart *ge?* can co-occur with reduplication to express basically the same aktionsart, leading to a doubly diminutive effect. Analyzing the aktionsart *ge?* as an Asp head would leave no hierarchical position for the reduplication morpheme RED in this situation.

For the above reasons, we do not analyze the aktionsart *ge?* as an aspectual head. We analyze it as a diminutive verbal modifier instead—that is, the modifier part of a modifier-head compound. Recall that this is just the original use scenario of Cat in Song (2019). We have already given the structure in (38c). See (41) for two more examples.

(41) a. *ge?-hua* 'scribble'



b. *ge?-shan* 'keep fanning (to make breeze)'



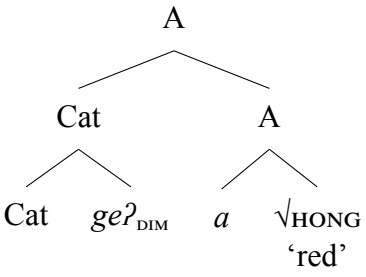
In particular, note that while the aktionsart meaning of *ge?* here is regular and basically compositional, the exact meaning of the overall *ge?*-prefixed verb is still up to lexicalization to some extent. For instance, *ge?-hua* does not just mean any iterated and diminutivized drawing action but more specifically means a casual scribbling action (e.g., on the wall), and one cannot use it for situations where, say, a painter is repeatedly applying small brushstrokes to an artwork. Such subtle idiomaticity in the meanings of *ge?*-prefixed verbs, despite their overall compositionality, lends support to our compounding analysis.

Cross-linguistically, aspectual verbal prefixes like *ge?* are not rare. For instance, there is a salient group of such elements in Hungarian, which are similarly called "verbal modifiers" in É. Kiss (2002) and described as intuitively "form[ing] a compound with the verb" (p. 56). See (42) for some example words with the telicizing verbal prefix *meg-*, which denotes that the verbal event necessarily reaches a natural endpoint.

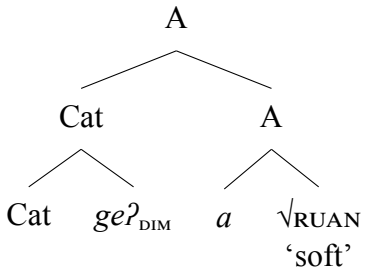
(42) *meg-eszik* 'MEG-eat; eat up', *meg-ír* 'MEG-write; write down' [Hungarian]
meg-talál 'MEG-find; find (telic)', *meg-hízik* 'MEG-put.on.weight; get fat (telic)'

Thus, our analysis of the aktionsart prefix *ge?* in Jin Chinese as a verbal modifier has cross-linguistic precedence. In particular, note that the telicizing function of *meg-*, albeit abstract and quasi-functional, is still considered a lexical semantic feature of *meg-* rather than a formal feature on some functional head, because *meg*-prefixed verbs are also still idiomatically lexicalized despite their basic compositionality, just like *ge?*-prefixed verbs are. As such, we similarly treat the iterative-diminutive function of the aktionsart *ge?* as an abstract semantic feature [DIM] on *ge?*. That is, we assume two *ge?*s in the lexicon: a plain one and a diminutive one (and we sill posit a third one in §3.3). The former is used at the root level, and the latter, at the word level. Recall from Section 2.1 that in some Jin subvarieties the *ge?* in *ge?*-prefixed adjectives may regularly have an '-ish' reading. We analyze such words in the same way (43).

(43) a. *ge?-hong* 'reddish'



b. *ge?-ruan* 'softish'

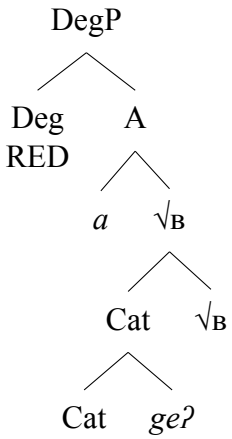


Note how the same diminutive *ge?* can yield different regular modifier meanings when combined with different categories: DIM + V = iterative-diminutive aktionsart, DIM + A = '-ish'. In this sense, the aktionsart *ge?* is not really category-specific either but still underlyingly cross-categorial: it is the manifestation of a generic $ge?_{\text{DIM}}$ in the verbal category.

3.3 Deriving the degree *ge?*

As for the degree infix *ge?*, recall from Section 2.3 that this use is the closest to an inflectional affix in Gaoping Jin. And to anticipate the content in Section 4, this use of *ge?* historically evolved under the joint influence of ABB and *ge?*-XX adjectives. In the previous section, we analyzed *ge?*-XX adjectives via a reduplicative Deg head on top of a *ge?*-prefixed adjective (38b). Can we analyze the *ge?*-BB part of an A-*ge?*-BB adjective in the same way (44)?

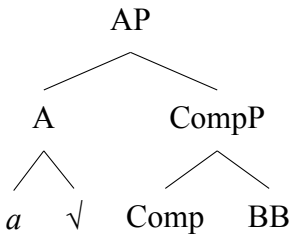
(44)



The answer is no, because *ge?*-XX adjectives are derived from *ge?*-X adjectives by reduplication, but the *ge?*-B part of an A-*ge?*-BB adjective normally does not exist as a word (e.g., *nan-ge-yiengyieng* 'blue-GE-BB' vs. **ge-yieng*)—nor do many *ge?*-BBs exist outside of the A-*ge?*-BB construction (e.g., **ge-yiengyieng*). Recall from Section 2.3 that A-*ge?*-BB adjectives in Gaoping Jin are formed by directly inserting *ge?* in existing ABB adjectives. Hence, an adequate analysis of them should start from a complete ABB form.

Partly following Z. Wang (2010) and Van Hoey (2023), we treat ABB adjectives as quasi compounds. That said, we do not derive them as words (i.e., big As) but treat them as phrasal idioms (APs). We will provide empirical evidence for this treatment below. For now, let us look at the structure in (45), which consists of a head A and a depictive complement BB, with the latter introduced by a null complement marker.

(45)

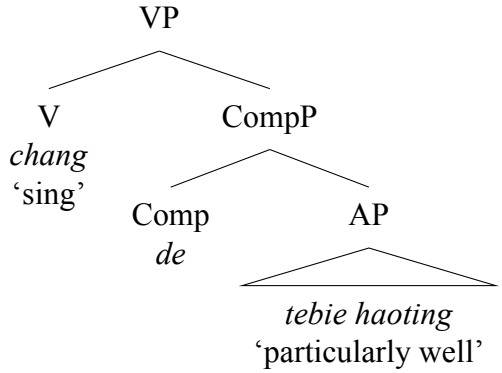


We call A a head in this structure, despite the root categorization representation, because we take the Comp-mediated combination of A and BB to be a derivational step in the main workspace, which operates on two prederived syntactic objects. We have used such cross-workspace derivation in Section 3.1 to derive *ge?*-prefixed classifiers; here we use it again. The assumption that syntactic derivation can—and indeed must—proceed in a multi-workspace or multilayered fashion is practically uncontroversial (though not often explicitly stated) in generative grammar (see, e.g., Zwart 2011; Fowlie 2013). Also recall from Section 2.3 that the A slot of an ABB adjective may be filled by a lexical verb or noun, as long as it is temporarily used adjectivally, denoting a state (see note 10). This is achieved in our proposed structure by the merger of a verbal/nominal root with *a*.[19] We abstract away from the internal structure of BB for current purposes; it could either be already in a reduplicated form in the lexicon or be reduplicated in syntax, or perhaps both options are available depending on the lexical nature of B.

Our choice to syntactically join A and BB via a complement marker follows a similar approach to manner/degree depictives in the verbal domain in Huang et al. (2009), where the complement marker is overtly realized as *de*, as in (46).

(46) a. *Ta chang de tebie haoting.* [Standard Mandarin]
he sing COMP particularly pleasant to hear
'He sings particularly well.'

b. (adapted from Huang et al. 2009: 87–89)



Intuitively, the BB part of an ABB adjective is a depictive complement for A just as the *de*-phrase in (46) is one for V. Indeed, we can often insert a *de* in ABB adjectives too, as in (47).[20]

(47) a. *Dalu leng de sousou de.* [Standard Mandarin]
continent cold COMP BB PRT
'It is quite chilly on the continent.' (ABB = *leng-sousou*)
(http://www.360doc.com/content/25/0320/15/32730899_1149462353.shtml)

[19] By "a verbal/nominal root" we mean a root that is frequently used in the verbal/nominal category and hence perceived as verbal/nominal by speakers. The formal-syntactic notion "root" is acategorial.

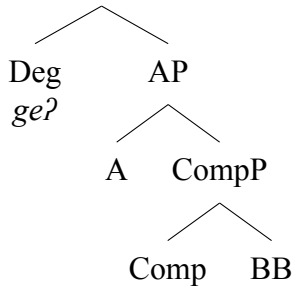
[20] These real-life examples are retrieved from the Internet on Aug. 6, 2025.

b. *Yiqie zhiwu dou lü de youyou de.*
every plant all green COMP BB PRT
'All plants are glossy and green.' (ABB = *lü-youyou*)
(https://czbooks.net/n/s6222k/s6gdg7kl?chapterNumber=1614)

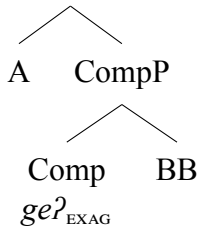
Thus, the possibility of *de*-insertion provides empirical evidence for our analysis of ABBs as phrasal idioms.

Having motivated our treatment of ABB adjectives, we now turn to A-*ge?*-BB. Our ABB structure in (45) makes available two possible analyses. On the one hand, we may merge *ge?* above the entire ABB structure as a degree head, as in (48a). In this analysis, the A-*ge?*-BB word order is derived by A-to-Deg movement. On the other hand, we may also merge *ge?* in Comp on a par with *de*, as in (48b). In this analysis, we can directly get the A-*ge?*-BB word order, without movement.[21]

(48) a.



b.



As far as we are concerned, both analyses may work, though in the second analysis (48b) we must encode the exaggerative meaning directly in *ge?* (via an abstract semantic feature [EXAG]) rather than in a syntactic position like Deg. There is no such restriction in the first analysis (48a). Here we are assuming that *ge?* spells out an exaggeratively valued Deg (i.e., Deg_{EXAG}), but as we will see in Section 4, the alternative possibility of encoding the exaggerative function in *ge?* is also useful from a diachronic perspective.

Semantically, applying the exaggerative function (EXAG) to ABB as a whole (48a) and to BB alone (48b) have no clear difference in effect, as in (49).

(49) a. ABB: *nan-yiengyieng* 'blue-BB; brightly blue'
b. *ge?*(ABB) = EXAG(brightly blue) = so [brightly blue]!!
c. A-*ge?*(BB) = EXAG(brightly) blue = [so brightly] blue!!

That said, one potential difference between the two structures in (48) lies in their influence on productivity. In (48a), since *ge?* merges above the entire ABB, it does not break the continuity of ABB in the underlying structure or at the C-I interface. By contrast, in (48b), *ge?* breaks the continuity of ABB at all levels of representation. Recall from Section 2.3 that the selectional relation between A and BB is idiomatic—so is the meaning of each ABB item (see also Van Hoey 2023). As such, the discontinuous syntactic representation in (48b) may affect the successful retrieval of an ABB's (stored) meaning at the C-I interface (in a DM model). In fact, the only way to guarantee successful meaning retrieval in (48b) is by lexicalizing the entire A-*ge?*-BB

[21] As to why BB does not move to Comp, we assume that this is because Comp is neither a categorizer for BB nor a functional category in its extended projection. The head movement cases we have seen so far all fall in one of these two scenarios.

chunk (i.e., giving it an entry in List 3 in DM). This means that (48b), though yielding the same surface string and semantic effect as (48a) does, is less productive. That may turn out to be a desirable result, because as mentioned in Section 2.3, the A-*ge?*-BB construction is not equally productive in different Jin subvarieties. We may assume that (48a) is active in subvarieties with productive A-*ge?*-BB, while (48b) is active in subvarieties where the construction is not (yet) productive.

Of course, this distinction in productivity relies on the assumption that the *ge?* in (48b) is not merely an exponent of Comp but exists in the underlying syntactic structure, presumably as an adjunct to Comp.[22] This is a reasonable assumption because, on the one hand, the exponent of Comp in Modern Chinese is *de* (or some dialectal counterpart of it) as mentioned above, and on the other hand, if *ge?* were an exponent of Comp in (48b), the Comp in question would have to carry an [EXAG] feature; namely, there would have to be two flavors of Comp in the lexicon—or even more than two flavors considering *ge?* still has two allomorphs (*be?* and *hue?*) with a negative evaluative effect (see §2.4). Such semantic flavoring is highly unusual and also unprecedented for a complement marker. As a final remark, recall that while the degree infix *ge?* is productive in Gaoping Jin, its negative allomorphs *be?* and (especially) *hue?* are not; this is also generally true across Jin subvarieties (see, e.g. Xing 2002). Thus, perhaps (48b) is available in Gaoping Jin too for sporadic A-*be?*/*hue?*-BB items, where the *be?*/*hue?* in Comp enter into an idiomatic selectional relation with BB.

4 The evolution of *ge?*

In this section, we extend the formal analysis in Section 3 to the diachronic aspects of *ge?*, showing that it can help us model the peculiar grammaticalization scenario observed in Section 1 in a principled way. We first review the history of *ge?* (§4.1), then examine its evolutionary pathway (§4.2), and finally formally represent its evolution in our model (§4.3).

4.1 A consequence of disyllabification

A mainstream view on the origin of *ge?* is that it had emerged as a consequence of the disyllabification tendency in the history of Chinese (see, i.a., Y. Liu 2001; Xing 2002; Y. Bai 2005). The lexicon of Old Chinese was predominantly monosyllabic, but due to the disyllabification tendency (roughly from Late Old Chinese onward), disyllabic words gradually became the new norm. Multiple disyllabification strategies were at work amid this huge transition, and what gave birth to *ge?* was a phonologically based operation called "syllable-splitting."[23] This works by separating the onset and rime of a syllable σ and creating out of them two new syllables σ_1 and σ_2 by filling in the missing parts—typically by inserting *-e?* [əʔ] as the rime of σ_1, and *l-* [l], as the onset of σ_2.[24] See (50) for some Jin Chinese words formed in this way.

[22] One way to implement this idea is via the generalized categorization operation we have used in §3.1 to derive *ge?*-prefixed classifiers; namely, deriving [$_{\text{Comp}}$ Comp *ge?*$_{\text{EXAG}}$] in a separate workspace, with *ge?*$_{\text{EXAG}}$ playing the role of a root, and then merging this "root-supported head" (Song's 2019 term) into the main workspace.

[23] See Lü (1989) and H. Wang (1999) for more phonological and morphological disyllabification strategies.

[24] These filler phonemes are used because, with [əʔ] being very weak and short and [l] being the most sonorous onset in Chinese, they together create a stark contrast in sonority between σ_1 and σ_2 (H. Wang 1999).

(50) a. *geng* 'low earth bank between fields' → *ge?-leng* [Shahe Jin]

b. *ba?* 'turn (something) around' → *be?-la?* [Taiyuan Jin]
huae 'ring, circle' → *hue?-luae* (H. Wang 1999: 190; X. Fan 2012: 157)

As we can see, in split-syllable words, σ_1 is not necessarily *ge?* [kəʔ] but may also be *be?* [pəʔ] and *hue?* [xuəʔ] among others depending on the onset of σ. Recall from Section 2.4 that *be?*/*hue?* have remained allomorphs of *ge?* to this day.

The disyllabification tendency itself was caused by the significant simplification of syllabic structure from Old to Middle Chinese (e.g., consonant clusters were lost, final stops became unreleased; Feng 2000b: 104; Zhengzhang 2003: 38), which gave rise to a large number of homophones and made communication difficult. Disyllabification served as a mechanism to re-establish lexical distinctions (Lü 1989; L. Wang 2004, 2005; Shi 2023).[25]

One argument in the literature for the syllable-origin hypothesis of *ge?* is that the geographical distribution of split-syllable words and that of *ge?*-words synchronically overlap, which is just the Jin-speaking area and adjacent regions (L. Li 2002; H. Fan 2007). Xing (2002) further notes that while both types of words thrived during the Song and Yuan dynasties (10th–14th centuries), the thriving of split-syllable words predated that of *ge?*-words. This implies that the former were a precursor of the latter. Besides, as Ma (1995) and Xing (2002) point out, all principled attempts to find the lexical origin of *ge?* have failed, which has led researchers to conclude that it had probably originated as a nonlexical element. In Xing's (2002: 277) words, the prefix *ge?* was born via reanalysis of the first syllable *ge?* in split-syllable words by analogy amid the "powerful current" of disyllabification.

As to why it was *ge?* (as well as *be?*/*hue?*) that acquired morphemic status, but not other σ_1s, Xing (2002: 277) assumes that this was due to frequency—of all the σ_1s, *ge?* was the most frequent and impressive one. In fact, it is still the most frequent σ_1 in split-syllable words in use today (see also X. Fan 2012: 159). And as to why *ge?*-words only systematically exist in the Jin-speaking area today, X. Liu (1995, 2008) attributes this to the geographical isolation of this area—it is located in the core zone of the Loess Plateau, which is characterized by high mountains and steep ridges (see also Y. Bai 2005). For this reason, X. Liu calls Jin Chinese the "mountain talk" of Northern Chinese. Qiao (2003) similarly remarks that Jin Chinese, though surrounded by Mandarin Chinese, has changed much more slowly due to its closed geographical environment.

Thus, the real situation is probably not that *ge?*-words only systematically exist in the Jin-speaking area but more exactly that they have only systematically survived there. Historically, *ge?*-words had apparently been a more widespread phenomenon in China. This is clear from Middle Chinese examples, which are not geographically restricted. For instance, the sentences in (51)—partly based on Xing (2002) and Y. Bai (2005)—are all from texts written outside of the Jin-speaking area. For each historical example below, we specify the writer, the time of writing (by dynasty and century), and the location of writing (using the name of the corresponding modern-day province).[26]

[25] An alternative hypothesis is that the simplification of syllabic structure left many monosyllabic words unable to make up a metrical foot on their own, which led to a reshaping of the foot makeup in Chinese from bimoraic to disyllabic (Feng 2000a, 2000b).

[26] We adopt Baxter and Sagart's (2014) romanization for Middle Chinese (H = rising tone, X = falling tone).

(51) a. *KaengH drjak ʼjit sraewng bje-gjaek-tsiX,* ***kət-thej*** [Middle Chinese]
further wear one pair leather-clog-N GE-ti.ONOM

kət-thap *tsyhwit mwon dzen.*
GE-tap.ONOM exit door front

'Moreover, (she) is wearing a pair of leather clogs and—clatter, clatter—comes out of the door.' (Onom) (Ya Cui, Tang, 9th c., Jiangsu)

b. *Mij-nye* ***kət-tsrjuwH****, hjweH dzywij mju ngjoX.*
eyebrow-N GE-furrow for whom not have words

'(My) brows furrow; for whom (am I) at a loss for words?' (V)
(Pang Mao, Song, 11th c., Zhejiang)

c. *Tsyin dzyeX* ***kət-tsong-nye****.*
really is GE-hate-R

'(He) is really annoying.' (A) (Juren Lü, Song, 11th/12th c., Anhui)

d. *Dzi-tsyewH yiX syuwX tsyijX dijH hjun: Bjaeng dijH khiX* ***kət-twoj****.*
Ci-Zhao with hand point ground say flat ground rise GE-pile

Zjwe tsyijX yowngX ʼjit twoj thuX.
follow point surge one pile earth

'Ci-Zhao (a Chan master) pointed to the ground with his hand and said: Let a mound of earth (i.e., a grave mound) rise from the flat ground. And a mound of earth rose following his pointing.' (N) (Dao Qian, Song, 12th c., Fujian)

These four sentences were respectively written in what are Jiangsu, Zhejiang, Anhui, and Fujian Provinces today, all of which are distant from Shanxi (the main Jin-speaking area). Note that all the *geʔ*-words in these historical examples have the word-formation prefix *geʔ*. This is evidently the oldest and historically most salient use of *geʔ*. It is also the use that is mainly (if not exclusively) focused on in historical discussions about *geʔ* in the literature. Also note that the *geʔ*-words in the above early examples are already cross-categorial, being used in N, V, A,[27] and Onom, just not in Cl. The absence of *geʔ*-prefixed classifiers from Middle Chinese is presumably due to the fact that classifiers themselves only became a standard part of Chinese syntax later in history, in the 14th century according to Shi (2023).

4.2 Trans-domain recycling

In the previous section, we saw that Jin Chinese *geʔ* had begun its life in history as a meaningless syllable—the first syllable of some split-syllable words—from where it got reanalyzed as a word-formation prefix. According to Xing (1987, 2002), analogy played a major role in this process. Besides, apparently a noticeable portion of the earliest *geʔ*-words (with either the syllable or the prefix *geʔ*) happened to denote "small" concepts, which was how *geʔ* came to be further associated with diminutivity.

The above reanalysis process is not a prototypical scenario of grammaticalization, if it could be called grammaticalization at all. Recall from Section 1 that grammaticalization is the diachronic process whereby a lexical item or construction takes on grammatical characteristics or through which a grammatical item becomes more grammatical (Hopper & Traugott 2003).

[27] The *-nye* in the adjective *kət-tsong-nye* (51c) corresponds to the nominal ending *-er* in Modern Chinese (see §2.1), though here it seems to be used as a generic derivational marker for disambiguation purposes—hence our convenient gloss R—as *kət-tsong* by itself could also be a verb.

Eventually, a highly grammaticalized item may become an integral part of a word, being downgraded from a morpheme to a mere phonological constituent (e.g., English *let's* → *lets*). In the case of *ge?*, however, it is a phonological constituent (a syllable) that has changed into a morpheme. Compared to grammaticalization, this looks more like *recycling*—the reuse of existing linguistic material for new purposes.

Grammatical recycling has long been observed across languages. For example, genitive case markers in Germanic languages tend to be recycled as linking elements in compound words (see Booij 2008 and references therein), as illustrated by the Dutch examples in (52).

(52) *schaap-s-kudde* 'sheep-LK-flock; flock of sheep' [Dutch]
her-en-huis 'gentleman-LK-house; mansion'
voorbehoed-s-middel 'prevent-LK-means; contraceptive'
witt-e-brood 'write-LK-bread; white bread' (Booij 2019: 146–147)

As Booij (2019) remarks, the use of originally nominal case endings on verbal and adjectival stems (e.g., *voorbehoed*, *witt*) is clear evidence that these morphemes have been repurposed.

Another example, discussed in Lass (1990), is the recycling of the aspect-coding ablaut pattern *e*∼*o*∼∅ from Proto-Indo-European as number indicators in past tense inflection in Germanic. See (53) for an illustration with examples from Ancient Greek and Gothic, with the former preserving the old system and the latter showing the new system.

(53)

Proto-Indo-European	**Present**	**Perfect**	**Aorist**
	e	*o*	∅
Ancient Greek	*le**í**p-ō*	*lé-l**o**ip-a*	*é-lip-on*
	'I leave'	'I have left'	'I left'
Germanic	**Present**	**Preterite$_{SG}$**	**Preterite$_{PL}$**
Gothic	*b**e**it-an*	*b**a**it* (*ai* < *oi*)	*bit-um*
	'to bite'	'I/he/she/it bit'	'we bit'

(Lass 1990: 84–85)

According to Lass, when the old perfect/aorist aspects got replaced by a single preterite tense, the remnants of the old ablaut system became "historical junk" and were "pressed into the service of new linguistic function" (p. 87). Lass calls this redeployment process "linguistic exaptation," borrowing a term from evolutionary biology: "[T]hose useful structures that arose for other reasons, or for no conventional reasons at all, and were then fortuitously available for other changes, we call exaptations." (Gould 1983: 171, via Lass 1990: 80).

We note that the syllable-to-morpheme change in the case of Jin Chinese *ge?* becomes completely normal when viewed from this broader perspective of recycling. Hence, we deem such a domain-general conception of recycling useful for our task at hand (though it need not be tied to "historical junk"). Such a broader perspective is also promoted by Biberauer (2017), who proposes that linguistic recycling reflects a general cognitive factor Maximize Minimal Means (in the context of Chomsky's 2005 "third factor" principles).

Although the initial, "bootstrapping" step in the evolution of *ge?* is not grammaticalization, the ensuing steps apparently are (though still atypical as we will see below). And while it is unclear to us when the two further uses of *ge?* (i.e., the aktionsart prefix use and the degree infix use) emerged in history, judging by their rarely being mentioned by scholars in historical discussions, we can tentatively assume that they are late developments in Jin Chinese. Indeed, their grammaticalization may still be ongoing today, considering their varied productivity in different Jin subvarieties. Xing (2002: 280) proposes the evolutionary pathway for *ge?* in (54).

(54) syllable *ge?* → prefix *ge?* → infix *ge?*

This pathway does not include the aktionsart use of *ge?* (or perhaps it is merged into the word-formation prefix use) but does include the degree infix use. Xing hypothesizes that the second step of development here (i.e., prefix → infix) was driven by Jin speakers' need to create four-syllable idioms—which are a most beloved form of idiom in Chinese since ancient times (Lü 1989: 207). Specifically, this took place under the joint influence of ABB adjectives and *ge?*-XX state-denoting words (including adjectives and onomatopoeic words). Recall from Section 2.1 that *ge?*-XX words are derived from *ge?*-prefixed words (i.e., *ge?*-X) by reduplication.

If Xing's proposal is on the right track, this further step of evolution still involves recycling and is not purely (classical) grammaticalization. Specifically, Jin speakers recycled the already actively used word-formation *ge?* as a filler syllable for ABB adjectives to batch-create four-syllable idioms. Then, this filler syllable *ge?* (as well as its allomorphs) got associated with a degree-marking function (i.e., exaggerative), partly due to the intrinsically vivid and expressive nature of ABB adjectives (see §2.3) and partly due to the melodic effect of the four-syllable idiom pattern (H. Wang 1999: 198). Thus, we can expand (54) to (55) by breaking down its second step.

(55) syllable *ge?* → prefix *ge?* → syllable *ge?* → infix *ge?*

This is clearly not unidirectional grammaticalization, though the second cycle of syllable-to-morpheme reanalysis, with an existing morpheme as its precursor, is no longer "out of thin air." The questions is: How can we formally model this diachronic process? We will present a solution in the next section based on our synchronic analysis of the multifunctionality of *ge?*.

4.3 Toward a formal pathway

First, recall from Section 3.2 that we did not analyze the aktionsart *ge?* as a category-specific function but analyzed it—as well as the '-ish' *ge?* in adjectives—as a continuation of the category-general word-formation *ge?* qua a root, only with an additional abstract semantic feature [DIM]. As such, the syllable *ge?* and the prefix *ge?* (in all its uses) can stay in the same grammaticalization chain. What chains them up is their shared vehicle of syntactic entrance, Cat. However, the degree infix *ge?* cannot be put in the same chain, as it does not enter syntax via Cat. Thus, we must split (55) into two branches.

Second, formally speaking, both the word-formation *ge?* and the aktionsart/'-ish' *ge?* involve the compounding use of a categorially underspecified morpheme *ge?*, either plain or diminutive, and recall from Section 3.3 that we have posited an exaggerative *ge?* as well. All these "flavors" of *ge?* qua a morpheme ultimately evolved from the syllable *ge?*, and in Section 3 we have strategically identified them all as roots, following the underspecification approach to root-augmenting elements promoted in X. Wang (2022) and De Belder et al. (2014). Now, we are ready to further motivate that identification from a diachronic perspective. When a piece of phonological material Φ is recycled for morphosyntactic use, it needs to be added to the lexicon. The "least effort" method is by directly turning it into the phonological index of a placeholder root—namely, a new lexicon entry ⟨PHON: i_Φ, SEM: ∅, F: ∅⟩ that has no other information beyond this phonological index.[28] We take this syllable-to-root reanalysis to be the default mechanism of

[28]We assume that each lexicon entry has three templatic slots—respectively for phonological, semantic, and formal information—and use i_Φ to indicate a phonological index that maps to Φ at PF. We also note that while such a "phonological root" is atypical under the usual conception of roots, it nevertheless meets the formal definition of a root and is furthermore compatible the conception in Borer (2013).

phonology-to-syntax recycling; it provides a threshold for further and more substantial changes to happen, which would require adding further information to the root and/or linking the phonological index to some functional category as its exponent. Thus, we posit that when the syllable /geʔ/ was first recycled for word-formation purposes (i.e., when it first acquired morphosyntactic status), it was added to the lexicon as a root whose sole (idiosyncratic) information was the phonological index for /geʔ/. We have been writing this root as *geʔ* in our discussion for expository convenience and also to distinguish it from normal roots like √DUEI 'pile', which additionally have idiosyncratic semantic content. Now we can define the root *geʔ* in a formally precise manner, as in (56).

(56) $ge\text{ʔ} := \langle \textsc{phon}\text{: } i_{/ge\text{ʔ}/}, \textsc{sem}\text{: } \emptyset, \textsc{f}\text{: } \emptyset \rangle$

Based on the above elaboration, we update the evolutionary pathway in (55) as (57). We distinguish the two syllable *geʔ*s by indices because they have different sources: the source of syllable $ge\text{ʔ}_1$ is split-syllable words, while that of syllable $ge\text{ʔ}_2$ is *geʔ*-prefixed words.

(57) syllable $ge\text{ʔ}_1 \longrightarrow$ **root *geʔ* (word-formation)** $\longrightarrow$ root $ge\text{ʔ}_{\textsc{dim}}$ (Akt, '-ish')

$\downarrow$

syllable $ge\text{ʔ}_2 \longrightarrow$ root $ge\text{ʔ}_{\textsc{exag}} \longrightarrow \textsc{Deg}_{\textsc{exag}}$ *geʔ* (A-*geʔ*-BB)

This is a highly unusual scenario of grammaticalization, even with the extensions of polygrammaticalization (Craig 1991) and transcategoriality (Robert 2018), as introduced in Section 1. First, the back-and-forth material-recycling between phonology and morphosyntax is a key feature in the evolution of *geʔ*. In this sense, Jin speakers are really making maximal use of minimal means (to reuse Biberauer's 2017 words). Second, the evolution of *geʔ* mainly involves root-level changes rather than changes at the level of functional categories. There is only one functional-category-level change: the emergence of an exaggerative Deg category. This matches our empirical observation that the multifunctionality of *geʔ* is largely confined to the lexical-derivational domain. Recall from Section 2 that the only use of *geʔ* close to an inflectional affix is the degree infix *geʔ*.

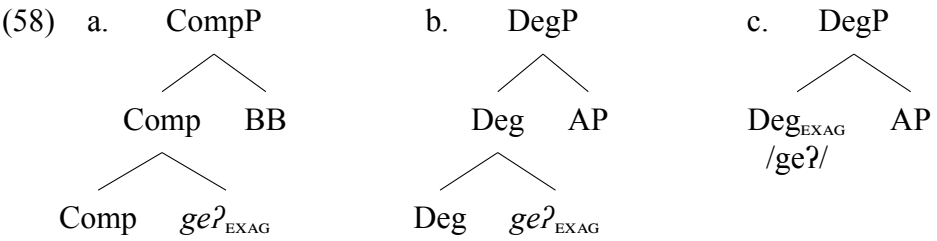
We further note that the root-level changes in (57) are accompanied by corresponding changes in their "entry points" to syntax. The plain root *geʔ*, the diminutive root $ge\text{ʔ}_{\textsc{dim}}$, and the exaggerative root $ge\text{ʔ}_{\textsc{exag}}$ respectively enter the derivation at the root level, the word level, and the level of phrasal syntax (via Comp). Two theoretical tools adopted in our model have made such flexible root usage possible: the defective categorizer (Cat) and generalized root syntax (both borrowed from Song 2019). Recall from Section 3.1 that Cat is an underspecified categorizer that works in tandem with a normal categorizer (via Agree and labeling) to derive compounding structures. On the other hand, generalized root syntax frees root usage from the traditional lexical domain and makes roots usable in the functional domain as well—to create root-supported or semifunctional (aka semilexical) heads. Root categorization and root support are two sides of the same coin: When a root and a categorizer merge, the root gets categorized and licensed in syntax, while the categorizer (defined as any functional category à la Song) gets supported by idiosyncratic content.

The theoretical advantage of the flexible entry points for roots is especially clear in the evolution of the degree infix *geʔ*. Recall from Section 4.2 that A-*geʔ*-BB adjectives historically emerged under the joint influence of ABB adjectives (vividly augmentative and expressive) and *geʔ*-XX state-denoting words (augmentative). As such, we can assume that the exaggerative effect in question was initially tied to the A-*geʔ*-BB construction instead of *geʔ*—and that it

was only reanalyzed as a function of *ge˨* later. The former stage does not rely on the existence of a functional head Deg_{EXAG}, whereas the latter does. Importantly, the postulation of such a new head—or a new value [EXAG] of an existing head Deg—is only possible when the A-*ge˨*-BB construction has gained some momentum; it would not suffice to have only a few sporadic items. Based on the above reasoning, we posit three more detailed stages in the evolution of the degree infix *ge˨*: *i*) an early stage without Deg_{EXAG}, *ii*) a transitional stage with a root-supported $\text{Deg}_{ge˨_{\text{EXAG}}}$, and *iii*) a final stage with a fully regular Deg_{EXAG} (pronounced /ge˨/).

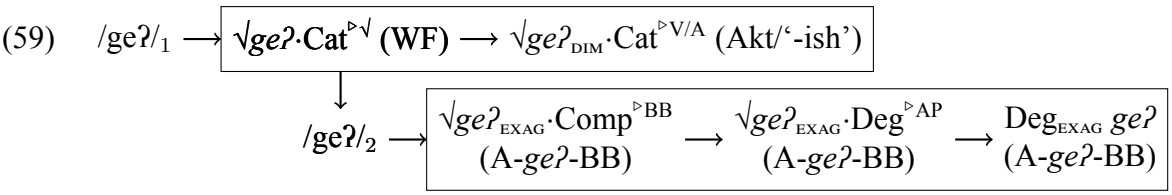
At stage (i), since there is no Deg_{EXAG} or formal feature [DEG: EXAG] yet, the reanalysis of the exaggerative function from the A-*ge˨*-BB construction to *ge˨* can only take place by putting a semantic feature on the latter, creating the root $ge˨_{\text{EXAG}}$. Also because there is no Deg_{EXAG} yet, this new root must join syntax in a different, makeshift way. We use the Comp head from Section 3.3 for this purpose. That is, we take the structure from (48b)—partly repeated in (58a)—to be a reflection of stage (i). As suggested in Section 3.3, this is the stage where the A-*ge˨*-BB construction is still a nonproductive phenomenon relying on the lexicalization of individual items. Next, at stage (ii), this use of *ge˨* has been identified as functional degree-marking, but the identification is not yet fixed as a dedicated Deg_{EXAG}. Instead, it is syntactically implemented as the root-support of an existing Deg head with a compatible value (e.g., [AUG(MENTED)]) via the exaggerative *ge˨* root, as in (58b). Finally, stage (iii) is the endpoint of the Deg_{EXAG} branch in (57), having a functional head with /ge˨/ as exponent. This stage is active in Jin subvarieties where the A-*ge˨*-BB construction is highly productive (e.g., Gaoping Jin). See (58) for a side-by-side comparison of the three stages.

(58)



Incorporating this further elaboration, we arrive at the final version of our proposed formal evolutionary pathway for *ge˨* in (59). We have abbreviated the node labels due to space limitations: the notation $\sqrt{\cdot}\text{X}^{\triangleright\text{Y}}$ (U) means that in the use context U the root $\sqrt{}$ enters syntax by supporting (i.e., being categorized by) the functional category X at point Y (which is the node that $\text{X}_{\sqrt{}}$ merges with), so $\sqrt{ge˨}\cdot\text{Cat}^{\triangleright\sqrt{}}$ (WF) means that in the WF (= word-formation) use the root *ge˨* enters syntax via Cat at the root level, and so forth.

(59)



All but one uses of the morpheme *ge˨* shown in this composite pathway coexist in Jin Chinese today—with the exception of the $\sqrt{ge˨}_{\text{EXAG}}\cdot\text{Deg}^{\triangleright\text{AP}}$ use for A-*ge˨*-BB adjectives since it is just a transitional stage. If we only look at the boxed subparts of (59), we see a scenario of generic transcategoriality in Robert's (2018) terminology, and if we further focus our attention on only either one of the two boxes, we see locally oriented transcategoriality (i.e., classical grammaticalization). Hence, the peculiar grammaticalization scenario of *ge˨* is essentially peculiar due

to its flexible mixture of grammaticalization and trans-domain recycling, with the latter gluing multiple instances of the former together. In other words, we observe a case of trans-domain polygrammaticalization via recycling, as the title of this paper indicates.

5 Conclusion

In this paper, we have studied the multifunctional affix *ge?* in Jin Chinese, which presents a peculiar scenario of grammaticalization. While classical, prototypical grammaticalization involves unidirectional changes between meaningful linguistic units along a cline of grammaticality (from less to more grammatical), the evolution of *ge?* is neither unidirectional nor entirely between meaningful linguistic units. Instead, it involves back-and-forth trans-domain changes between phonology and morphosyntax. Neither standard grammaticalization theory nor its polygrammaticalization or transcategoriality extension can satisfactorily explain this situation, so we have resorted to a broader, domain-general notion of recycling—or linguistic exaptation (Lass 1990). This notion may reflect a general cognitive factor (aka "third factor") Maximize Minimal Means (Biberauer 2017).

We have characterized the evolution of *ge?* as trans-domain polygrammaticalization via recycling. To support this characterization, we have carefully studied both the synchronic multifunctionality and the diachronic development of *ge?*. Its three synchronic functions observed in the entire Jin-speaking area are a cross-categorial word-formation prefix, an iterative-diminutive aktionsart verbal prefix, and an exaggerative degree infix for ABB adjectives. We have analyzed the word-formation prefix *ge?* as an "auxiliary" root introduced by a defective categorizer Cat (Song 2019) to augment another root; the aktionsart prefix *ge?*, as a verb-modifying root with a diminutive semantic feature, also introduced by Cat; and the degree infix *ge?*, as either a functional head Deg_{EXAG} or a root with an exaggerative feature adjoined to an abstract complement marker in the ABB structure, depending on whether the use is productive in a Jin subvariety. Diachronically, the word-formation use of *ge?* emerged early as a consequence of the impactful disyllabification tendency in the history of Chinese, while the aktionsart and the degree-marking use of *ge?* are presumably late or even ongoing developments. The former is a direct continuation of the word-formation use (by grammaticalization), and the latter is a creative reuse of the prefix *ge?* (by exaptation). The *ge?*-phenomenon has only systematically survived in the Jin-speaking area today because of its isolated geographical environment in premodern times.

The affix *ge?* has received much previous research in the literature on Jin Chinese, but previous works on it are almost all descriptive and published in Chinese. By systematically introducing this "locally famous" morpheme to the international linguistic community and formally modeling it within an up-to-date theoretical framework, we hope to contribute to both the synchronic study of multifunctionality and the diachronic study of grammaticalization.

References

Acedo-Matellán, V., & Real-Puigdollers, C. (2019). Roots into functional nodes: Exploring locality and semi-lexicality. *The Linguistic Review*, *36*(3), 411–436.

Bai, P. (1988). Ge fei citou bian [*Ge?* is not a prefix]. *Journal of Shanxi University (Philosophy and Social Science Edition)*, *2*, 60–63.

Bai, Y. (2005). Jinyu ge zi yanjiu [On *ge?* in Jin Chinese]. *Linguistic Research*, *1*, 57–59.

Baxter, W., & Sagart, L. (2014). *Old Chinese: A new reconstruction*. Oxford University Press.

Biberauer, T. (2017). Factors 2 and 3: A principled approach. In C. Song & J. Baker (Eds.), *Cambridge occasional papers in linguistics* (pp. 38–65, Vol. 10). Theoretical; Applied Linguistics, University of Cambridge.

Booij, G. (2008). *Recycling morphology: Case endings as markers of Dutch constructions* [Manuscript, Leiden University].

Booij, G. (2019). *The morphology of Dutch* (2nd ed.). Oxford University Press.

Borer, H. (2013). *Taking form* (Vol. 3). Oxford University Press.

Cavirani-Pots, C. (2020). *Roots in progress: Semi-lexicality in the Dutch and Afrikaans verbal domain* [Doctoral dissertation, KU Leuven].

Chomsky, N. (1986). *Barriers*. MIT press.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, *36*(1), 1–22.

Chomsky, N. (2013). Problems of projection. *Lingua*, *130*, 33–49.

Chomsky, N. (2015). Problems of projection: Extensions. In E. D. Domenico, C. Hamann, & S. Matteini (Eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti* (pp. 1–16). John Benjamins.

Craig, C. (1991). Ways to go in Rama: A case study in polygrammaticalization. In B. Heine & E. C. Traugott (Eds.), *Approaches to grammaticalization* (pp. 455–492, Vol. 2). John Benjamins.

De Belder, M., Faust, N., & Lampitelli, N. (2014). On a low and a high diminutive: Evidence from Italian and Hebrew. In A. Alexiadou, H. Borer, & F. Schäfer (Eds.), *The syntax of roots and the roots of syntax* (pp. 149–163). Oxford University Press.

Dong, X. (2021). Shanggu hanyu zhong de chongdie ji qi yanbian: ABB shi xingrongci cifa moshi de lishi xingcheng [Reduplication in Old Chinese and its change]. *Lingnan Journal of Chinese Studies*, *15*, 269–290.

É. Kiss, K. (2002). *The syntax of Hungarian*. Cambridge University Press.

Fan, H. (2007). *Dingxiang fangyan yufa yanjiu* [A study on the grammar of Dingxiang Dialect]. Language & Culture Press.

Fan, X. (2012). Cong shahehua de geleici kan hanyu yizhong xingtai chengfen de xingcheng [A study of a morphological element's formation in Chinese based on the investigation of *ge?*-words in Shahe Dialect]. *Language Sciences*, *11*(2), 157–172.

Feng, S. (2000a). Hanyu shuangyinhua de lishi laiyuan [The historical origin of disyllabification in Chinese]. *Xiandai zhongguoyu yanjiu [Contemporary Research in Modern Chinese]*, 123–138.

Feng, S. (2000b). *Hanyu yunlü jufaxue* [Chinese prosody syntax]. Shanghai Educational Press.

Fortin, A., & Rainer, F. (2022). Evaluative suffixes. In A. Ledgeway & M. Maiden (Eds.), *The Cambridge handbook of Romance linguistics* (pp. 434–458). Cambridge University Press.

Fowlie, M. (2013). Multiple multiple spellout. In T. Biberauer (Ed.), *Challenges to linearization* (pp. 129–169). Mouton de Gruyter.

Gaeta, L. (2013). Affix ordering and conversion: Looking for the place of zero. *Lingue e linguaggio*, 145–170.

Gould, S. J. (1983). *Hen's teeth and horse's toes: Further reflections in natural history*. Norton.

Halle, M., & Marantz, A. (1993). Distributed Morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *Essays in linguistics in honor of Sylvain Bromberger* (pp. 111–176). MIT Press.

Hopper, P., & Traugott, E. (2003). *Grammaticalization* (2nd ed.). Cambridge University Press.

Hou, J. (1999). *Xiandai jinyu de yanjiu* [Studies on Modern Jin Chinese]. Commercial Press.

Huang, C.-T. J., Li, Y.-H. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge University Press.
Huo, L. (2015). Qianxi shanxi pingding fangyan zhong de ge [A preliminary study on *ge?* in Pingding Dialect in Shanxi Province]. *Modern Chinese*, (10), 37–39.
Kayne, R. (1994). *The antisymmetry of syntax*. MIT Press.
Kiefer, F. (1996). Prefix reduplication in Hungarian. *Acta Linguistica Hungarica*, *43*(1/2), 175–194.
Kiefer, F. (2010). Areal-typological aspects of word-formation: The case of aktionsart-formation in German, Hungarian, Slavic, Baltic, Romani and Yiddish. In F. Rainer, W. U. Dressler, D. Kastovsky, & H. C. Luschützky (Eds.), *Variation and change in morphology* (pp. 129–148). John Benjamins.
Lass, R. (1990). How to do things with junk: Exaptation in language evolution. *Journal of Linguistics*, *26*(1), 79–102.
Li, L. (2002). Fangyan bijiao, juyu fangyan shi yu fangyan fenqu: Yi jinyu fenyinci he fuzhou qiejiaoci wei li [Dialectal comparison, history of dialectal areas and dialectal division: A case study of *fenyinci* in Jin dialects and *qiejiaoci* in Min dialects]. *Fangyan [Dialect]*, (1), 41–59.
Li, S. (2024). *Cross-categorial multifunctionality from a distributed morphological perspective: The case of ge- in Gaoping Jin* [Master's thesis, Zhejiang University].
Liu, C.-S. L. (2013). Reduplication of adjectives in Chinese: A default state. *Journal of East Asian Linguistics*, *22*, 101–132.
Liu, X. (1995). Zailun hanyu beifanghua de fenqu [Revisiting the classification of northern Chinese dialects]. *Zhongguo Yuwen [Chinese Language]*, *6*, 447–454.
Liu, X. (2008). Huangtu gaoyuan de fangyan shi yige baozang [The dialects of the Loess Plateau are a treasure trove]. *Yuwen Yanjiu [Studies in the Chinese Language]*, *2*, 37–41.
Liu, Y. (2001). Jinyu cihui shuangyinhua de yizhong fangshi: Jia ge [A method of disyllabification in the vocabulary of Jin Chinese: Adding *ge?*]. *Zhongguo Yuwen [Chinese Language]*, *1*, 83–84.
Lü, S. (1989). Xiandai hanyu danshuang yinjie wenti chutan [A peliminary inquiry into the issue of mono- and disyllabic words in Modern Chinese]. In *Lü Shuxiang zixuanji [Lü Shuxiang's self-selected anthology]* (pp. 181–210). Shanghai Education Publishing House.
Ma, W. (1995). Jin fangyan li de ge zi [*Ge?* in Jin dialect]. *Journal of Shanxi Datong University*, *3*, 38–43.
May, R. (1985). *Logical form: Its structure and derivation*. MIT press.
Potts, C. (2007). The expressive dimension. *Theoretical Linguistics*, *33*(2), 165–198.
Qiao, Q. (2003). Jinyu yu guanhua fei tongbu fazhan 1 [The non-synchronous evolution of Jin Chinese and Mandarin Chinese (part 1)]. *Fangyan [Dialect]*, *2*, 147–160.
Qiao, Q. (Ed.). (2023). *Shanxi fangyan yufa yanjiu* [Studies on Shanxi dialect grammar]. China Social Sciences Press.
Robert, S. (2018). The challenge of polygrammaticalization for linguistic theory. *Cognitive Linguistic Studies*, *5*(1), 106–132.
Roberts, I., & Roussou, A. (2003). *Syntactic change: A minimalist approach to grammaticalization*. Cambridge University Press.
Shao, J. (1990). ABB shi xingrongci dongtai yanjiu [Dynamicity in ABB adjectives]. *Shijie Hanyu Jiaoxue*, *11*, 19–26.
Shi, Y. (2023). *The evolution of Chinese grammar*. Cambridge University Press.
Song, C. (2019). *On the formal flexibility of syntactic categories* [Doctoral dissertation, University of Cambridge].

Song, C. (2025). Adjunction as categorization: On the syntactic quirkiness of word-level modification. In H. Kishimoto, M. Kawashima, & Z. Xu (Eds.), *Papers from the International Workshop on the Syntax of Predication and Modification 2024* (pp. 139–158).
Tian, J. (2015). Shanxi xiaoyi fangyan zhong de getouci [*Ge?*-prefixed words in Xiaoyi Dialect of Shanxi Province]. *Xue Zhou Kan*, (5), 20–21.
Van Hoey, T. (2023). ABB, a salient prototype of collocate–ideophone constructions in Mandarin Chinese. *Cognitive Linguistics*, *34*(1), 133–163.
Wang, H. (1999). *Hanyu feixianxing yinxihue* [Nonlinear phonology of Chinese]. Peking University Press.
Wang, L. 王力. (2004). *Hanyu shigao* (2nd ed.) [A draft history of the Chinese language]. Zhonghua Book Company.
Wang, L. 王力. (2005). *Hanyu yufa shi* [A history of Chinese grammar]. The Commercial Press.
Wang, L. 王利. (2017). *Jin dongnan jinyu lishi bijiao yanjiu* [A comparative study on the history of Southeastern Jin Chinese]. China Social Sciences Press.
Wang, L. (2001). Shanxi fangyan getouci de jiegou leixing [Structural types of *ge*-prefixed words in Shanxi dialects]. *Studies of the Chinese Language*, *1*, 80–82.
Wang, X. (2022). *Ciji qudong de jinyu gouci yu yinxi jiaohu moshi* [Word-based interaction of morphology and phonology in Jin Chinese]. Commercial Press.
Wang, Z. (2010). The head of the Chinese adjectives and ABB reduplication. In L. E. Clemens & C.-M. L. Liu (Eds.), *Proceedings of the 22nd North American Conference on Chinese Linguistics (NACCL-22) & the 18th International Conference on Chinese Linguistics (IACL-18)* (pp. 232–245, Vol. 1).
Xing, X. (1987). Jinyu getouci liubianlun [A study of the evolution of *ge?*-prefixed words in Jin Chinese]. *Journal of Inner Mongolia Teachers University (Philosophy and Social Science Chinese Edition)*, *2*, 78–85.
Xing, X. (2002). *Shenmu fangyan yanjiu* [Studies on Shenmu dialect]. Zhonghua Book Company.
Yang, J. (2008). *Hanyu fangyan xingrongci chongdie yanjiu [A study of adjectival reduplication in Chinese dialects]* [Doctoral dissertation, Fudan University].
Yin, G. (2016). *Shanxi fangyan chongdie yanjiu [A study on reduplication of Shanxi dialects]* [Master's thesis, Shanxi University].
Zhang, L. (2020). Gezici 60 nian yanjiu gaishu [An overview of research on *ge?*-words in the past 60 years]. *Journal of Xinzhou Teachers University*, *36*(4), 34–39, 62.
Zhang, T. (2021). Shanxi Gaoping fangyan ge-V jiegou fenxi [A study on the structure of *ge*-V in Gaoping Dialect in Shanxi Province]. *Journal of Jincheng Institute of Technology*, *14*(5), 5–8.
Zhengzhang, S. (2003). *Old Chinese phonology*. Shanghai Educational Publishing House.
Zwart, J.-W. (2011). Recursion in language: A layered-derivation approach. *Biolinguistics*, *5*(1–2), 43–56.