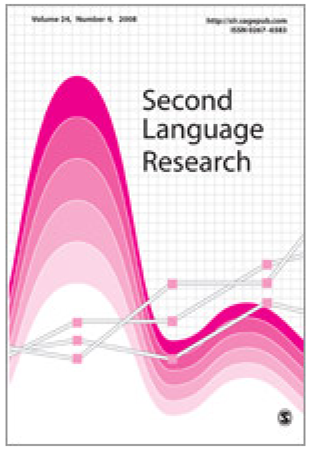

Dutch-Mandarin learners' on-line use of syntactic cues to anticipate mass vs count interpretations

Journal:	*Second Language Research*
Manuscript ID	SLR-22-0065.R3
Manuscript Type:	Original Article
Keywords:	Implicit learning, Mass/count syntax, Second language acquisition, Unique-to-L2 construction, Visual World Paradigm
Abstract:	It remains unclear whether late second language learners (L2ers) can acquire sufficient knowledge about unique-to-L2 constructions through implicit learning to build anticipations during real time processing. To tackle this question, we conducted a Visual World Paradigm experiment to investigate high-proficiency late L1-Dutch L2-Mandarin Chinese learners' on-line processing of syntactic cues to count vs. mass interpretations in Chinese which are unique-to-L2 and never explicitly taught. The results showed that late Dutch-Mandarin learners were sensitive to a mass-biased syntactic cue in real time processing, and exhibited some native-like anticipatory behaviour. These findings indicate that late L2ers can acquire unique-to-L2 constructions through implicit learning, and can automatically use this knowledge to make predictions.

Dutch-Mandarin learners' on-line use of syntactic cues to anticipate mass vs count interpretations

Abstract: It remains unclear whether late second language learners (L2ers) can acquire sufficient knowledge about unique-to-L2 constructions through implicit learning to build anticipations during real time processing. To tackle this question, we conducted a Visual World Paradigm experiment to investigate high-proficiency late L1-Dutch L2-Mandarin Chinese learners' on-line processing of syntactic cues to count vs. mass interpretations in Chinese which are unique-to-L2 and never explicitly taught. The results showed that late Dutch-Mandarin learners were sensitive to a mass-biased syntactic cue in real time processing, and exhibited some native-like anticipatory behaviour. These findings indicate that late L2ers can acquire unique-to-L2 constructions through implicit learning, and can automatically use this knowledge to make predictions.

Keywords: Implicit learning, Mass/count syntax, Second language acquisition, Unique-to-L2 construction, Visual World Paradigm

Introduction

In the field of L2 acquisition, unique-to-L2 constructions play an important role in revealing insights about the learning and processing mechanisms of late L2ers which could not be related to transfer effects from L1. Due to the salience of unique-to-L2 constructions, many are highlighted and taught explicitly in language learning classes to L2ers. This makes it difficult to investigate whether unique-to-L2 constructions can be implicitly acquired. Compared to explicit learning scenarios in which learners are instructed to actively look for patterns intentionally, implicit learning refers to a process in which learners derive knowledge from a complex, rule governed stimulus domain without intending to and without becoming aware of the knowledge they have acquired (Reber, 1967). Native speakers acquire their L1 mostly through implicit learning as children. Late L2ers, however, acquire most of their L2 knowledge through explicit learning (in classrooms, with explicit and specific instructions and corrections, etc.), though implicit learning may also occur in some circumstances. In order to get a full picture of L2 acquisition, it is important to explore L2ers' implicit learning, especially of unique-to-L2 constructions. Through a Visual World Paradigm experiment, the current study aimed to investigate whether L2ers can acquire unique-to-L2 constructions through implicit learning, and use this knowledge to build anticipations during real time processing. To be specific, the critical question of the current study is whether high-proficiency L1-Dutch L2-Mandarin learners (Dutch-Mandarin learners hereinafter, the 1st and 2nd languages of L2ers are represented in the linear order of the terminology of L1-L2 learners correspondingly) can predictively use mass/count syntactic cues which are unique-to-Mandarin constructions and can only be implicitly acquired.

Implicit learning in L2 acquisition

Implicit learning has been observed in artificial materials learning (Leung & Williams, 2014; Rebuschat & Williams, 2012) and L2 acquisition (Cleary & Langle, 2007; Robinson, 2005). Some researchers found that unique-to-L2 constructions can be implicitly learned by late L2ers (Donaldson, 2011; Montrul & Slabakova, 2003; Slabakova, 2006; Tolentino & Tokowicz, 2014). Donaldson (2011) found that near-native English-French learners demonstrated native-like behaviours in the production of left dislocation, which is a linguistic construction that does not exist in English and is rarely explicitly taught in language classrooms. English-French learners can only pick it up through implicit learning in an immersive French environment. Montrul & Slabakova (2003) investigated advanced English-Spanish learners' use of Preterit and Imperfect past tenses, which exist in Spanish but not English. By using a truth value judgment task, they found that very advanced English-Spanish learners can eventually exhibit native-like behaviour in spite of the poverty of the stimulus. These findings indicate that even after puberty, late L2ers can still implicitly learn unique-to-L2 constructions.

However, it should be noted that previous studies exploring L2ers' implicit learning often use off-line tests or training & post-test paradigms, which restrict language learners' learning time and experience: participants only have limited time to learn the hidden grammar in very unnatural conditions. It is therefore important to explore late L2ers' implicit learning over a long period in an immersive L2 environment and under natural linguistic input conditions. Moreover, it remains unclear whether unique-to-L2 constructions which can only be implicitly acquired can be automatically used by L2ers in real time processing.

Anticipation building in L2 processing

Anticipation building plays an important role in real time sentence processing. Native speakers often proactively integrate different sources of information (e.g. lexical, semantic, syntactic, etc.) rapidly, and combine these inputs to build an evolving representation and possible anticipations for the upcoming items (Altmann & Mirković, 2009; Federmeier et al., 2007; Huettig et al., 2011; Kamide, 2008; Pickering & Garrod, 2013). For L2ers, however, there is no consensus about their predictive processing capacities, let alone their predictive use of unique-to-L2 constructions (Grüter et al., 2012; Grüter & Rohde, 2013 vs. Hopp, 2012, 2015; Trenkic et al., 2014).

Grüter et al. (2012) investigated whether highly proficient English-Spanish learners can use the gender information on determiners to facilitate the interpretation of upcoming nouns. Gender-marking is unique-to-Spanish knowledge for English-Spanish learners. In their experiment 3 (a looking-while-listening experiment), there were eight determiner-noun pairs in the audio stimuli, four of which were masculine and four feminine. The visual stimuli were presented in two trial types: same-gender trials where the two images depicted objects referred to by nouns of the same grammatical gender; and different-gender trials where images depicted objects referred to by nouns of different genders. The authors predicted that if participants can use gender-marking on determiners to anticipate the interpretation of the upcoming nouns, faster shifting to the target image on different-gender trials than on same-gender trials would be expected. They found that native Spanish speakers were faster to converge on the target picture on different-gender trials compared to same-gender trials, while the difference was not significant for English-Spanish learners. This finding indicated that unlike native speakers who can use gender information on the determiner as a predictive cue in online processing, English-Spanish learners (even with a high level of proficiency) were less effective in anticipatory use of this unique-to-L2 construction. Based on these observations, Grüter & Rohde (2013) proposed that non-native speakers have reduced ability to generate expectations during language processing.

However, other studies revealed native-like patterns for intermediate to high proficiency L2ers in predictive use of unique-to-L2 constructions. Hopp (2012) tested advanced to near-native English-German learners' predictive processing of syntactic gender agreement between determiners and nouns in German in a Visual World Paradigm experiment. Due to the lack of gender marking in English, the gender agreement between determiners and nouns is unique-to-German knowledge to English-German learners. In the experiment, participants were required to look at four pictures presented on the screen, and at the same time listen to audio recordings from the headphones. They had to choose one out of the four pictures as a target based on the auditory stimuli. Each auditory sentence had the structure of 'Where is **the**$_{masc/fem/neut}$ yellow [noun]?', where the gender status of the determiner was manipulated to have three levels: masculine, feminine, and neuter. An adjective which is unmarked for gender and can occur with all three gender levels of nouns was used. Focusing on which picture participants start to fixate on when hearing the gender-marked determiner, Hopp et al assessed whether participants used the gender information from the determiner to predict the upcoming gender-consistent noun. The results showed that near-native English-German learners, but not advanced ones, exhibited native-like patterns in the predictive use of the gender information from the determiners: they fixated on the pictures that denoted the gender-consistent nouns on hearing the determiners, well before they heard the nouns. This finding indicates that with increased L2 proficiency, late L2ers can ultimately use unique-to-L2 constructions to build anticipations. Similar findings were reported in Hopp (2015) who found that high-proficiency English-German learners can automatically use case markers (which are unique-to-German for English-German learners) in anticipation building.

Such native-like behaviour in real-time anticipation-building has also been found in intermediate L2ers. Trenkic et al. (2014) conducted a Visual World Paradigm experiment to test Mandarin-English learners' on-line comprehension of English articles. Articles are a unique-to-L2 structure for Mandarin-English learners since Mandarin Chinese is an article-less language. The visual materials were manipulated to have two conditions: a two-compatible referent condition and a one-compatible referent condition. The definiteness status of the target nominal phrase was manipulated in the audio materials ('The [agent] will put the [theme] inside **the**/**a** [goal]'). For example, the audio materials could be either 'The pirate will put the cube inside **the** can' or 'The pirate will put the cube inside **a** can', while the corresponding visual materials would be either a picture where there are **two** opened cans, or a picture where there is only **one** opened can. Participants were required to look at the pictures on the screen and at the same time, listen to the descriptions about what is going to happen in the picture from the headphones, and then to mouse-click on the location on the screen where the described object will end up when they finish listening. The authors expected that if late Mandarin-English learners can take advantage of articles in English, they should start to land more fixations on the target faster in the audio-visual matched conditions (i.e. definite article ***the*** + one-compatible referent picture; indefinite article ***a*** + two-compatible referent picture) than in the audio-visual mismatched conditions (i.e. definite article ***the*** + two-compatible referent picture; indefinite article ***a*** + one-compatible referent picture). On the other hand, if late Mandarin-English learners cannot process articles on-line and rely overly on pragmatic cues, they should react faster when there is only one possible referent in the picture than when two referents are available, regardless of the definiteness status of the target nominal phrase in the audio stimuli. The results revealed a native-like pattern in intermediate Mandarin-English learners in anticipatory use of articles, although they were slower to make use of the article semantics than native English speakers.

In general, there are varying opinions about late L2ers' ability to make predictive use of unique-to-L2 constructions. It should be noted that the unique-to-L2 constructions used in previous studies can be learned explicitly. For example, constructions such as gender-marking on determiners in German and Spanish, and articles in English, all are highlighted and taught explicitly and specifically in language learning classes. Little is known about whether late

L2ers can predictively use unique-to-L2 constructions which can only be attained through implicit learning.

To fill this gap, the present study explored late Dutch-Mandarin learners' on-line processing of nominal phrases with different mass/count syntactic cues. For Dutch-Mandarin learners, the mass/count distinction marked in the classifier system and the mass/count-biased syntax structures are unique-to-L2 constructions given the absence of a classifier system in Dutch. Furthermore, mass/count syntactic cues are implicit to native Mandarin speakers as well as Dutch-Mandarin learners. Native Mandarin speakers automatically use them in daily life without being able to spell out the specific rules. Dutch-Mandarin learners are only taught about the obligatory appearance of a classifier between a numeral and a noun, and some high-frequency classifier-noun pairings, but not about the mass/count associations of different classifiers or about mass/count syntactic structures. They can only implicitly learn this knowledge through immersive experiences in a Mandarin-dominant environment, since no textbooks of Mandarin learning ever include this knowledge. Thus, mass/count syntactic structures in Mandarin offer us ideal natural language materials to investigate late language learners' implicit learning of unique-to-L2 constructions under natural conditions. To access Dutch-Mandarin learners' automatic use of mass/count syntactic structures, the present study conducted a Visual World Paradigm experiment which can offer us a fine-grained index of participants' attention shifts evoked by their interpretation/understanding of the linguistic input during real time processing.

Syntactic mass/count cues in Chinese

The mass/count distinction is a universal concept. Regardless of what language one speaks, one has the world knowledge that objects can be roughly divided into two kinds: objects with discrete units and clear boundaries, which can be counted based on number (e.g., cats); and substances without specific units or fixed boundaries, which can only be measured based on volume (e.g., water) (Barner et al., 2009; Choi et al., 2017; Gleitman & Papafragou, 2005; Li et al., 2009). Though the ontological concepts of mass/count objects/substances are independent of specific languages, the mass/count distinction is instantiated in different ways in different languages (Borer, 2005; Cheng & Sybesma, 1998, 1999, 2012; Cheng et al., 2008; Chierchia, 1998a, 1998b, 2010; Doetjes, 1997).

In Dutch, the mass/count distinction is reflected in overt singular and plural markings. The indefinite article *'een'* is used to mark singularity (1a), while plural morphemes *'-en/-s'* are used to mark plurality (1b). To make a mass noun countable, a measure word has to be inserted between the numeral and the noun (1c), and it is the measure word which undergoes the inflectional plural change (1d).

(1) a. ***een*** *boek*
a book
'a book'

b. *drie boek**en***
three book**s**
'three books'

c. *een **stuk** kaas*
a **piece** cheese
'a piece of cheese'

d. *twee stukk**en** kass*
two piece**s** cheese
'two pieces of cheese'

Unlike Dutch, there is no number morphology to mark singularity or plurality on nouns[1], and neither a definite nor indefinite article in the nominal domain in Mandarin Chinese. Bare nouns in Chinese can be interpreted as either singular or plural, definite or indefinite given the appropriate contexts, without any morphological changes in the form (2a). A classifier is obligatory when combining a noun with a numeral (2b & 2c). Based on these observations,

[1] Some researchers claim that '*-men*' is a plural morpheme in Mandarin Chinese which can only occur with pronouns and nouns with a [+human] feature (Iljic, 1994; Choi et al., 2017). We put aside a discussion of '*-men*' in this study, and leave open the question of whether there is a plural morpheme in Mandarin Chinese comparable to that in English or Dutch. Critically, '*-men*' cannot occur with the inanimate nouns we investigate in this study.

researchers claim that the mass/count distinction in Chinese is marked in syntax and realized at the classifier level (Borer, 2005; Li et al., 2008; Yao et al., 2022).

(2) a. *wo kan jian* ***mao*** *le*
I look see **cat** LE
'I saw a cat/cats/the cat(s)'

b. *san* ***zhi*** *mao*
three **CL** cat
'three cats'

c. *liang* ***bei*** *shui*
two $\mathbf{CL_{cup}}$ water
'two cups of water'

Classifiers in Chinese can be roughly divided into two groups based on their function: count-classifiers, which simply name the unit that the entities inherently have; and massifiers, which create a unit to measure substances and pluralities (Allan, 1977; Tai, 1992; 1994; Tai & Wang, 1990). Count-classifiers differ from massifiers in several ways. First of all, count-classifiers do not have concrete lexical meanings, and merely play a grammatical role when a noun needs to be counted. Massifiers, on the other hand, can sometimes be nouns with concrete meanings. Secondly, count-classifiers differ from massifiers in the restrictions of classifier-noun pairings. Count-classifiers usually occur only with canonical count nouns (CCN) with countable singular readings (denoting individual objects with clear boundaries), while massifiers can occur with either canonical mass nouns (CMN) or with CCNs with plural/divided-portion readings. For example, the count-classifier *zhi* in (3a) can occur with *mao* 'cat' which (canonically) has discrete units, but not *you* 'oil' which does not. In (3b), *ping* 'bottle' is a massifier and can occur with *shui* 'water' which does not have specific units. Critically, though it can also occur with *ganlan* 'olive', which canonically is a count noun and expresses a plural reading in this specific case.

(3) a. *yi zhi mao /*you*
one CL cat/ *oil
'a cat/*an oil'

b. *yi ping shui/ ganlan*
one CL_{bottle} water/ olive
'a bottle of water/olives/*olive'

More relevant to the current study, there are two distributional differences between count-classifiers and massifiers (Cheng & Sybesma, 1998, 1999): only massifiers can be followed by the modification marker *de* (e.g., (4a) vs. (4b)), and only massifiers can be modified by adjectives like *da* 'big' and *xiao* 'small' (e.g., (5a) vs. (5b)). In other words, the structure [Num-Adj-Cl-(*de*)-Noun] (Cl is short for Classifier, N is short for Noun) could be treated as a massifier-biased structure (Li et al., 2008), and the structure [Num-Adj-Cl] could be treated as a syntactic massifier-biased cue which can be used automatically by native Mandarin speakers to make a mass/plural nominal interpretation in on-line processing (Yao et al., 2022). The differences between count-classifiers and massifiers are summarized in Table 1.

(4) a. *san* *bei* *(de)* *shui*
three CL$_{cup}$ DE water
'three cups of water'

b. *liang* *tou* *(*de)* *niu*
two CL$_{head}$ DE cow
'two cows'

(5) a. *san* *da* *zhang* *zhi*
three big CL$_{piece}$ paper
'three big pieces of paper'

b. **san* *da* *zhi* *gou*
three big CL dog
Intended reading: 'three big dogs'

Table 1
The differences between count-classifiers and massifiers

	Count-Classifier	**Massifier**
Semantic function	Naming units objects naturally have	Creating ways of measurement
Properties	Without concrete lexical meanings	With lexical meanings
Cl-N pairing	Canonical count nouns only	Both canonical count and canonical mass nouns
Insertion of *de*	×	√
[Adj-Cl] order	×	√

Previous studies tested the syntactic marked mass/count distinction in Chinese by using 'dual-role' classifiers and the massifier-biased syntactic cues. In Chinese, there exist some 'dual-role' classifiers which can be interpreted as either count-classifiers or massifiers given the appropriate context (see the discussion in Zhang, 2012). For example, the classifier *ba* in (6) can occur with both CCNs like *yaoshi* 'key' and CMNs like *shazi* 'sand'. *Ba* can be interpreted as a count-classifier when it occurs with a CCN such as *yaoshi* 'key'. In this case, it describes the unit a key naturally has. Thus the nominal phrase *yi ba yaoshi* can have the meaning of 'one key'. *Ba* can also be interpreted as a massifier, with either a CCN such as *yaoshi* 'key' or a CMN such as *shazi* 'sand'. In this case, it creates a way (handful) to measure the amount of keys or sand. Thus the nominal phrases *yi ba yaoshi/shazi* can have the meaning of 'a handful of keys/sand'.

(6) *yi* *ba* *jiandao/ yaoshi/ shaozi/ shazi/ shizi*
one CL$_{handful}$ scissors/ key/ spoon/ sand/ pebble
'one pair of scissors, a key/spoon, a handful of keys/spoons/sand/pebbles'

These dual-role classifiers are ambiguous between either count-classifier or massifier meanings. Different structures of nominal phrases can be treated as syntactic cues biased towards different meanings. To be specific, the phrase structure [Num-Adj-Cl-Noun] should indicate that the classifier in this case is a massifier. A neutral structure with no massifier-biased cue (e.g., [Num-Cl-Noun] or [Num-Cl-Adj-Noun]) would allow either a count-classifier or a massifier. Thus, putting dual-role classifiers in nominal phrases with different structures allows us to investigate how syntactic structures determine the mass/count interpretations of nominal phrases.

Previous studies found that native Mandarin speakers are sensitive to the massifier-biased syntactic structure. Li et al. (2008) conducted an off-line task with some dual-role classifiers and found that nominal phrases with the massifier-biased structure [Num-Adj-Cl-*de*-N] were always mapped to unshaped substances without specific units or objects organized in plural sets, while phrases with the neutral structure [Num-Cl-Adj-N] were mapped to objects with discrete units. Furthermore, in an eye-tracking Visual World Paradigm experiment (Yao et al., 2022), on hearing the [Num-Adj] word order, native Mandarin speakers immediately increased their rate of fixations to pictures depicting mass/plural-expressing meanings, indicating that they were expecting to hear a massifier, and were building a mass/plural interpretation accordingly. These findings indicate that the massifier-biased syntactic cue (i.e., the [Num-Adj-Cl] structure) can be used by native Mandarin speakers in both off-line and on-line processing.

As for Dutch-Mandarin learners, in order to learn the mass/count nominal expressions in Chinese, they need to acquire some unique-to-L2 knowledge: (1) the semantic and syntactic features of classifiers; (2) the restrictions on classifier-noun pairings; (3) the ambiguity of dual-role classifiers; (4) the massifier-biased syntactic cues. All of these exist in Chinese but not in Dutch (Lau & Grüter, 2015; Leung & Williams, 2014). While the first and second properties are explicitly taught in Chinese-learning classes, the third and fourth are not, and therefore can only be implicitly acquired. Native Mandarin speakers automatically use dual-role classifiers and the massifier-biased syntactic cues in daily life without being able to describe the specific rules. None of the textbooks regularly used to teach Chinese to L2ers introduces and discusses these facts. Even expert Chinese teachers are unaware of them when explicitly asked[2]. Dutch-Mandarin learners are only taught about the obligatory appearance of a classifier between a numeral and a noun, and some high-frequency classifier-noun pairings, but not about the mass/count associations of different classifiers or about any other features of the syntactic

[2] An on-line survey was completed by 20 Chinese teachers from the Confucius Institutes at UCL, Queen Mary University of London, SOAS University of London, and Leiden University. This survey explicitly asked these Chinese teachers about their awareness of the massifier/count-classifier distinction, and the syntactic cues that co-vary with this distinction. The results indicated that these Chinese-teachers have very little awareness of the count/mass distinction in the classifier system in Mandarin Chinese, even though they could use count-classifiers and massifiers automatically and unconsciously. They have no experience of teaching or encountering any instruction about count/mass-classifiers. They report that their Chinese-learning students have occasionally asked questions about the classifier-noun pairing restrictions, but none of them have ever mentioned the difference between count-classifiers and massifiers.

structures associated with count vs. mass interpretations.[3] They can only learn these features implicitly through immersive experience in a Chinese-dominant environment. As such, the dual-role classifiers and the massifier-biased syntactic cues offer us a great opportunity to explore late L2ers' implicit learning of unique-to-L2 knowledge, by investigating a real linguistic construction embedded in its natural linguistic context.

Previous studies used either off-line tests (Gong, 2010; Liang, 2008; Polio, 1994) or on-line tasks (Lau & Grüter, 2015) to explore L2-Chinese learners' understanding and processing of nominal phrases with different classifiers. They found that: (1) L2ers are sensitive to the obligatory occurrence of classifiers between a numeral and a noun, but not the semantic connections between classifiers and nouns (Liang, 2008; Polio, 1994); (2) L2ers can take advantage of classifier information during on-line processing to build anticipations for upcoming nouns (Grüter et al., 2020; Lau & Grüter, 2015). However, it remains unclear whether L2-Chinese learners can automatically use massifier-biased syntactic cues to build predictions in real time processing. Using a VWP experiment, the current study aimed to explore high proficiency late Dutch-Mandarin learners' on-line processing of nominal phrases with mass/count syntactic structures and dual-role classifiers, with a focus on their use of the [Num-Adj-Cl] structure to generate predictions. By comparing participants' behaviour in the current study to native Mandarin speakers' (Yao et al., 2022), we intended to tackle the question of whether late L2ers can predictively use unique-to-L2 constructions (which can only be acquired through implicit learning) in a native-like manner.

Method

[3] As based on the structure of Hanyu Shuiping Kaoshi since 2010, Official HSK Centre Introduction, HSK Centre; the Official Chinese teaching text book – Edexcel GCSE Chinese Student Book. The Mandarin Chinese classes are in English, as are the standard textbooks.

1. Materials

To allow for a direct comparison with native Mandarin speakers, the materials in the current study are the same as in Yao et al. (2022). For the audio stimuli, three dual-role classifiers were used: *gen* 'rod', *kuai* 'chunk', and *ba* 'handful'. Each classifier was embedded in a nominal phrase with three different structures: [Cl-N] is the **baseline** with the structure 'number-classifier-noun', [Cl-A-N] (A is short for Adjective) is the **neutral** condition for either count-classifier or massifier, and with the structure 'number-classifier-adjective-noun', and [A-Cl-N] is the **critical** massifier condition with the structure 'number-adjective-classifier-noun'. The number *san* 'three' was used in all audio materials. The adjectives *da* 'big' and *xiao* 'small' were used, with the units of objects/entities presented in the pictures manipulated to be either consistent or inconsistent with the adjectives in the audio sentences. Twelve CMNs and 12 CCNs were selected based on a Rating Test[4], in which CCNs were rated as nouns denoting objects which cannot be divided, while CMNs were rated as nouns denoting substances which can be divided, and would not lose their original features after dividing. Each noun contains two written characters; the number of syllables of each character and the lexical frequency of each noun were matched between CMNs and CCNs. Eight nouns occur with the classifier *ba* 'handful' (four CCNs & four CMNs), ten nouns occur with the classifier *kuai* 'chunk' (five CCNs & five CMNs), and six nouns occur with the classifier *gen* 'rod' (three CCNs & three CMNs). For each pair of nouns (a CCN and a CMN which share the same classifier), in order to avoid the *Phonological Competition Effect* (Klein et al., 2012) and *Tone Sandhi Effect* (Yip, 2002), the first character of each noun was controlled to have the same tone (and most of them share the same first syllable). Each of these nominal phrases (NPs) were embedded in sentences with the structure 'From the four pictures, could you please choose + [NP]'.

[4] Detailed information of this rating test is provided in the Appendix.

For the visual stimuli, there were four pictures for each pair of nouns. Two of them contained entities denoted by CMNs, and the other two contained objects denoted by CCNs. Based on the semantic difference between the classifier *ba* and the classifiers *gen* and *kuai* (*ba* is a collective classifier when interpreted as a massifier, while *gen* and *kuai* are dividers when interpreted as massifiers; see Cheng & Sybesma, 2012), entities on the pictures were organized in different ways. The detailed descriptions of visual stimuli for *gen*, *kuai* and *ba* are summarized in Figure 1 to Figure 3 respectively. In each figure, the red lines enclose the pictures intended to evoke a mass or plural interpretation (A, C & D), compatible with the interpretation of the sentences in the critical condition (with the structure [A-Cl-N]). Picture A contains objects compatible with the size adjective in the auditory sentence, picture D contains a size-incompatible object, and can thus be ruled out if adjective semantics are rapidly interpreted.

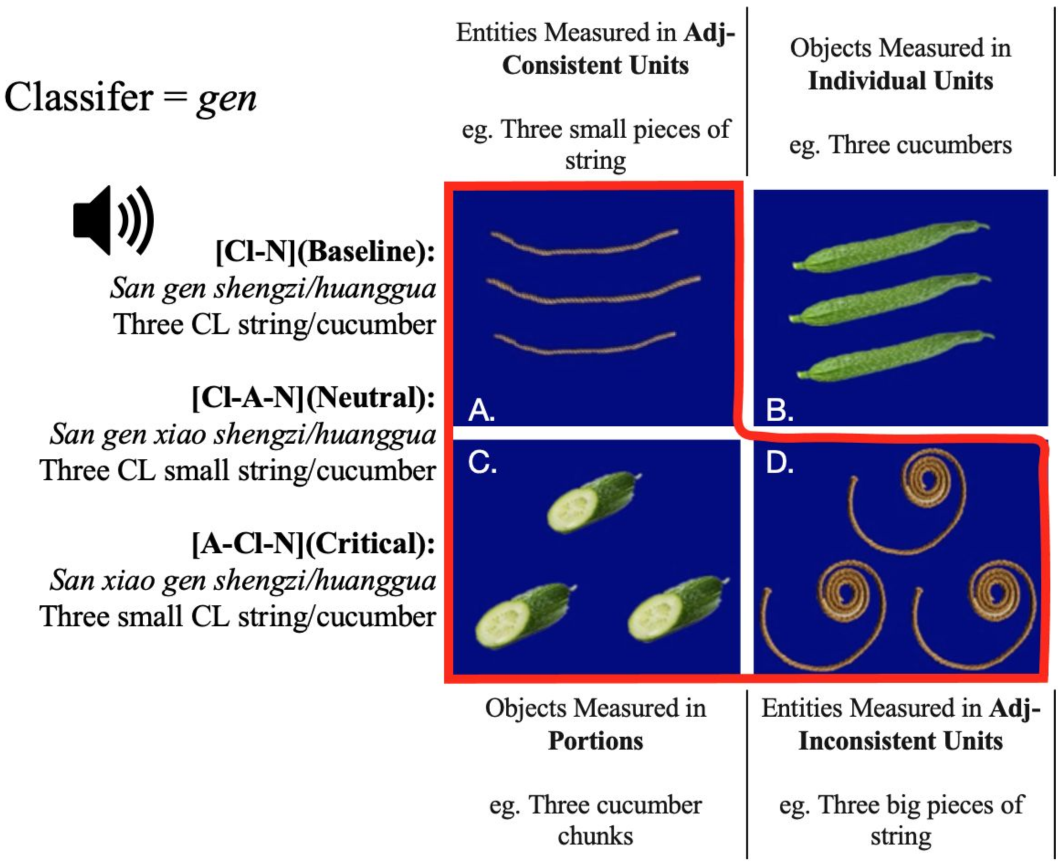

Figure 1

Example displays for classifier gen with example auditory materials.

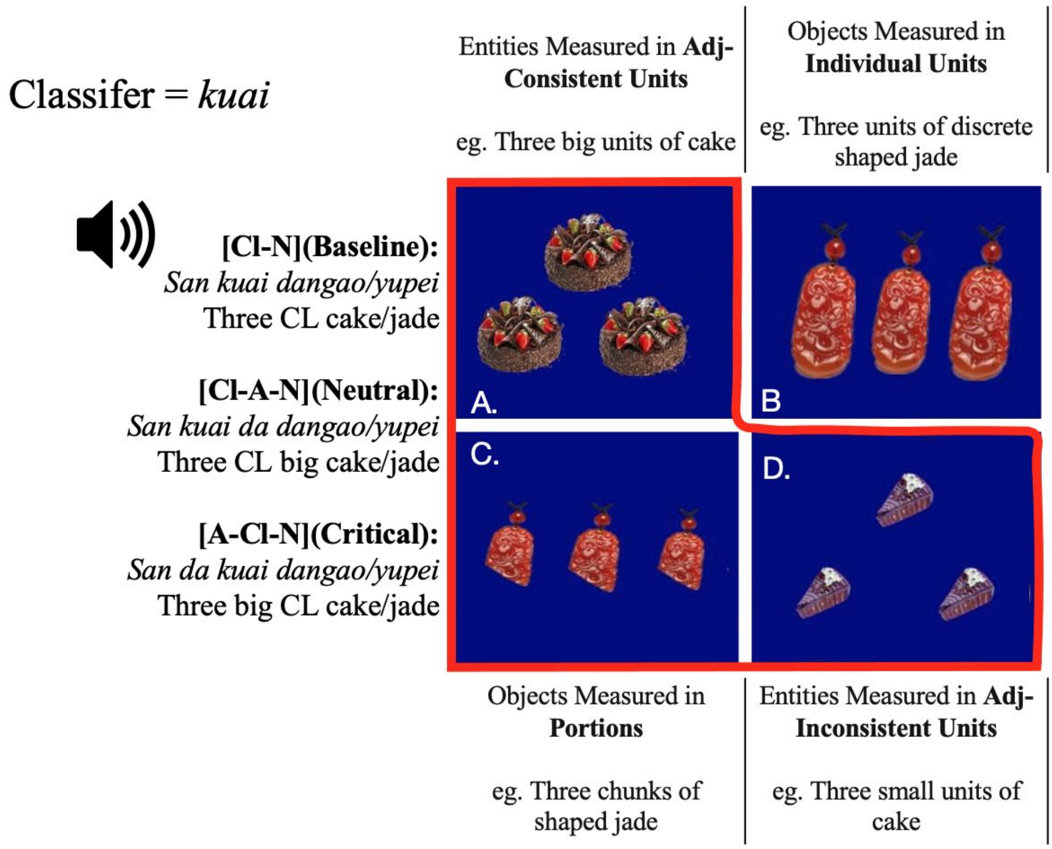

Figure 2

Example displays for classifier kuai with example auditory materials.

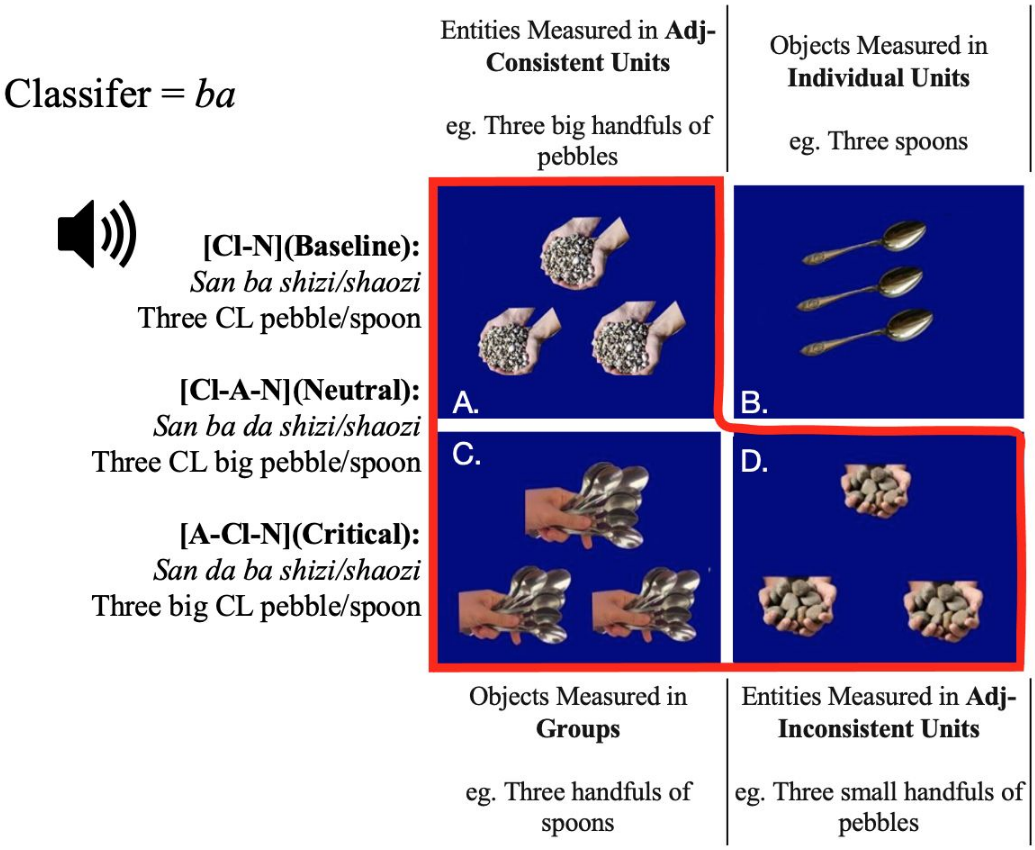

Figure 3
Example displays for classifier ba with example auditory materials.

In total, there were 72 audio sentences (24 nouns * 3 conditions). Each audio sentence was always accompanied with a set of four pictures (two of which contained the nouns in the sentence, while the other two contained the nouns paired with the nouns in the sentence). For example, when the audio sentence was *Qing cong si fu tu zhong xuanze* ***san da ba shaozi*** 'From the four pictures, could you please choose **three big handfuls of spoons**', and the noun paired with 'spoon' was 'pebble', then the four pictures simultaneously occurred with the audio sentence would be (A) a picture containing three big handfuls of pebbles – 'entities measured in Adj-consistent units', (B) a picture containing three individual spoons – 'objects measured in individual units', (C) a picture containing three handfuls of spoons – 'objects measured in groups', (D) a picture containing three small handfuls of pebbles – 'entities measured in Adj-inconsistent units' (as illustrated in Figure 3). These 72 audio sentences were pseudo-randomly divided into three lists to make sure that each list only contained one of the three conditions of each nominal. There were 24 fillers in each list which makes the total number of sentences in each list 48. The critical nominal phrases in the fillers do not contain any classifiers. These 144 sentences (72 critical audio materials + 72 fillers) were digitally recorded by a female native speaker of Mandarin Chinese in a sound-proofed booth, sampling at 44.1 KHz. Each participant was only tested on one of the three lists. The location of the four pictures in each trial was counterbalanced in each list. The whole experiment lasted around 30 minutes.

2. Participants

Thirty high proficiency late Dutch-Mandarin learners (16 females) participated. They were students from Leiden University majoring in Chinese (aged between 16 to 35 years old). They all started learning Mandarin Chinese after puberty, had been learning Chinese for at least 2 years and had experience of studying in mainland China for at least 6 months. They all passed the HSK-C (Hanyu Shuiping Kaoshi – advanced level, the standard Chinese language proficiency test for non-native speakers administered by the Ministry of Education of the People's Republic of China), which indicated that they were all high-proficiency L2-Chinese learners. Each participant was given €15 for their participation. The data was collected in Leiden University, Leiden, Netherlands.

3. Procedure

3.1 Naming Test

To make sure that all the pictures in the current study are recognizable and understandable to participants, before the critical VWP experiment, participants were asked to name all the pictures by writing down both Dutch and Chinese names, and to speak out loud the name of the objects in Mandarin Chinese. The accuracy of their written answers and pronunciations of the pictures were recorded and analysed. Only participants who were able to accurately name all objects and substances depicted in the critical materials were included as participants in the VWP experiment.

3.2 Visual World Paradigm Experiment

SR Research Eyelink 1000 was used to measure participants' eye movements. Both eyes' movements were recorded, but only the right eye's data were analysed. Following the nine-point calibration and validation, gaze-position error was less than 0.5°. Participants were tested in a sound-proof booth and seated 60cm from a 19-inch monitor.

Participants were tested individually. Before critical experimental trials, there were instructions and 10 practice trials. After participants read the instruction and finished the practice trials, a standard 9-point grid calibration and validation was completed. During the experiment, participants were asked to listen to the sentences through the headphones. At the same time, in each trial, there were four pictures presented on the screen. Participants were required to choose one out of the four pictures based on the sentences they heard by moving the mouse to click the corresponding picture. Participants' gaze was directed to the fixation cross in the middle of the screen prior to each trial to avoid baseline effects (Barr et al., 2011; Hopp, 2016). A trial only started when participants fixated on the calibration dot stably. Participants' eye movements during the display of the audio materials and their responses were recorded.

3.3 Fill in the Blanks Test

The Naming test ensured that L2 participants were familiar with all the nouns (both the written names and the pronunciations). To further make sure that they were familiar with the classifier-noun pairs used in the current study, a Fill in the Blanks test was conducted on the same group of participants after they completed the VWP experiment. In this test, each participant was presented with the numeral *yi* 'one' and a pair of parentheses preceding each noun, in a structure like [one ( ) Noun]. Participants were asked to fill in appropriate classifiers in the parentheses. They could also include multiple classifiers if they wanted to.

4. Predictions

As we are most interested in participants' predictive use of the mass/count syntactic structures, we focused on their fixation patterns occurring during each of the audio sentences before the nouns were heard. Before discussing predictions of Dutch-Mandarin learners' fixation patterns, we summarize the native Mandarin speakers' fixation patterns from Yao et al. (2022) in Table 2. In Yao et al. (2022), 30 native Mandarin speakers participated, all students from Beijing Normal University, aged between 18 to 28 years old. The data were collected in Beijing Normal University, Beijing, China.

Table 2
Summary of native Mandarin speakers' fixation patterns. A, B, C and D refer to the pictures in Figure 1-3.

	Baseline (Cl-N)	Neutral (Cl-A-N)	Critical (A-CL-N)
Gen & Kuai	CL evoked no preferences	Adj evoked more fixations on A and B	Adj evoked more fixations on A,C & D
Ba	More fixations on B, upon heading CL	Adj evoked no preferences	Adj evoked more fixations on A,C & D

Based on previous studies, there are three possible patterns high-proficiency Dutch-Mandarin learners could exhibit. First, L2ers may exhibit a native-like pattern in their predictive use of mass/count syntactic structures. Like native Mandarin speakers, L2ers could land more fixations on the mass/plural-expressing pictures than the other pictures on hearing the [A-Cl] structure (critical condition). This pattern would indicate that L2ers can implicitly acquire the mass and count meanings of dual-role classifiers and predictively use the structure [A-Cl] as a massifier-biased syntactic cue for interpretation. Alternatively, L2ers may be able to predictively use the mass/count syntactic structures to make the corresponding interpretations, but be slower than native speakers in using these cues on-line. In this scenario, L2ers could use the syntactic structure [A-Cl] to build a mass-biased interpretation for the whole NP, but may only exhibit their preference for mass/plural-denoting pictures at some point later than the adjectives (maybe on hearing the classifiers). Last but not least, L2ers may not be able to automatically use the mass/count syntactic cues during on-line processing at all. In this case, the [A-Cl] structure would be considered a neutral structure, compatible with both count and mass interpretations. This is because in Dutch, nominal phrases have structures such as [Num-Adj-Noun (+suffix)], [Num-Adj-Measure word-Noun], and [Num-Measure word-Adj-Noun]. The word order [Num-Adj] in Dutch does not offer any specific information about the mass/count status of the upcoming items. L2ers' interpretation of nominal phrases with the structure [A-Cl-N] should be determined by their understandings of the classifiers, which should be consistent with the dominant meanings of the classifiers in daily life. No difference would be expected between the [A-Cl-N] structure and the neutral structures [Cl-(A)-N].

5. Results

The results of the Naming test showed that all the pictures were recognizable and understandable to participants. They were familiar with the objects and substances presented in all the pictures, and were able to write down and speak out the right names.

The results of the Fill in the Blanks test showed that in most cases the first classifier that participants filled in for each noun was the classifier used in the current study. Even if for some nouns the first filled-in classifier was not the classifier used in current study, the second or third most frequently offered classifiers were. This result indicated that participants were familiar with the classifier-noun pairings used in the current study. Two out of 30 participants used the general classifier *ge* for more than a quarter of the nouns. Based on Polio (1994), this overgeneralized use of the classifier *ge* indicated a lack of full acquisition of the classifier-noun pairings. Hence, these two participants were excluded from the final analysis.

In analysing participants' eye movements in the VWP experiment, their fixation distributions among the four pictures along the display of audio nominal phrases and their final choice of pictures were reported. The former tells us how participants predictively process each item of the nominal phrases in real time and the latter reveals their final interpretation of the nominals. Furthermore, participants' data in the current study were compared to native speakers' data in Yao et al. (2022).

5.1 Fixation Proportions

Following Engelhardt & Ferreira (2010) and Tanenhaus et al. (1995) we used Linear Mixed Effects Modelling (LME; Baayen et al., 2008) for each classifier in each Region of Interest (ROI) to test how the fixation proportions to the four pictures changed in different nominal phrases. ROI is defined as the region extending from 200ms following the onset of a critical word (i.e., classifier/adjective/noun) to 200ms following the offset of this item. This 200ms buffer following the onset of a word is based on the mean time required to plan and launch an eye movement, and the typical lag observed between eye movements and fine-grained phonetic detail in the speech stream (Allopenna et al., 1998; Kukona et al., 2011). Fixation proportions to each picture in each ROI were calculated for 50ms time-bins. The on-line processing of the items occurring before the nouns should not be affected by the canonical mass/count status of the nouns so the fixation distributions in the ROI of Classifier and Adjective were analysed by collapsing across CCNs and CMNs. Fixation distributions in the ROI of Noun were analysed separately for CCNs and CMNs. Since the key aim of the current study is to investigate how fixation proportions in each picture changed on the occurrence of each critical item, we conducted priori custom contrasts to compare every two pictures successfully in each condition (Schad et al., 2020). Specifically, picture (the four types of visual display in three audio stimuli conditions) was included as a fixed effect in the model, with sliding contrasts (a contrast was made between every two neighbouring levels) applied to the four types of pictures in each condition. By doing so, we could test whether the occurrence of each audio item evoked a significant divergence of fixations. Statistical analysis was conducted in R (R Development Core Team, 2015). All the figures in the current study were generated with *ggplot2* (Wickham, 2009). Detailed results for each ROI are shown in the Appendix.

5.1.1. The Classifier region

The fixation distribution patterns for *ba*, *gen* and *kuai* were similar to each other in the classifier region. Thus the results for these three classifiers are reported together. With the baseline and neutral structures [Cl-N] and [Cl-A-N], fixations were randomly distributed among the four pictures: the fixation proportions were not significantly different from each other, *p*s > .1. With the critical structure [A-Cl-N], more fixations were landed on the mass/plural-denoting pictures than the singular/individual-denoting picture, *p*s < .001. The bar plots of the fixation proportions in the classifier window for *ba*, *gen* and *kuai* are illustrated in Figure 4-1A, 1B and 1C respectively.

5.1.2. The Adjective region

The fixation distribution patterns for *ba*, *gen* and *kuai* were similar to each other in the adjective region. With the neutral structure [Cl-A-N] and the critical structure [A-Cl-N], more fixations were landed on the 'entities measured in Adj-consistent units' pictures than the other three pictures, *p*s < .001. The difference between any two of these three pictures was not significant, *p*s > .1. The bar plots of the fixation proportions in the adjective window for *ba*, *gen* and *kuai* are illustrated in Figure 4-2A, 2B and 2C respectively.

5.1.3. The Noun region

5.1.3.1. Canonical Count nouns

(A). When the classifier was ba

With the baseline and neutral structures [Cl-N] and [Cl-A-N], on hearing CCNs, more fixations were landed on the CCN-denoting pictures compared to the CMN-denoting pictures, *p*s < .001. Additionally, the 'objects measured in individual units' picture attracted more fixations than the 'objects measured in groups' picture, *p*s < .001. On hearing the critical structure [A-Cl-N], the 'objects measured in groups' picture attracted significantly more fixations than the other three pictures, *p*s < .001. The bar plot of the fixation proportions for the classifier *ba* is illustrated in Figure 4-3A.

(B). When the classifier was gen and kuai

In all three conditions, participants exhibited the same fixation proportion patterns for *gen* and *kuai*: on hearing CCNs, more fixations were landed on the CCN-denoting pictures than the CMN-denoting pictures, *p*s < .001, and the 'objects measured in individual units' pictures attracted more fixations than the 'objects measured in portions' pictures, *p*s < .001. The bar plots of the fixation proportions with the classifier *gen* and *kuai* are illustrated in Figure 4-3B and 3C respectively.

5.1.3.2. Canonical Mass nouns

(A). When the classifier was ba

With the baseline structure [Cl-N], more fixations were landed in the CMN-denoting pictures than the CCN-denoting pictures, *p*s < .001, and with no significant difference between the CMN-denoting pictures. With the neutral structure [Cl-A-N], more fixations were landed in the CMN-denoting pictures than the CCN-denoting pictures, *p*s <. 001, but there were more fixations in the 'entities measured in Adj-inconsistent units' pictures than the 'entities measured in Adj-consistent units' picture, $p <. 001$. With the critical structure [A-Cl-N], more fixations were landed in the CMN-denoting pictures than CCN-denoting pictures, *p*s $<. 001$, and there were more fixations to the 'entities measured in Adj-consistent units' pictures than the 'entities measured in Adj-inconsistent units' pictures, $p <. 001$.The bar plot of the fixation proportions with classifier *ba* is illustrated in Figure 4-4A.

(B). When the classifier was gen and kuai

The fixation distribution patterns for *gen* and *kuai* were similar to each other and are reported together. With the baseline structure [Cl-N], on hearing the CMNs, more fixations were landed on the CMN-denoting pictures than the CCN-denoting pictures, *p*s $< .001$. There were more fixations on the 'entities measured in Adj-consistent units' pictures than the 'entities measured in Adj-inconsistent units' pictures with *gen*, $p < .05$, but the difference between these two CMN-denoting pictures was not significant with *kuai*, $p = .73$. With the neutral and critical structures [Cl-A-N] and [A-Cl-N], more fixations were landed on the CMN-denoting pictures than CCN-denoting pictures, *p*s $< .001$, and there were more fixations on the 'entities measured in Adj-consistent units' pictures than the 'entities measured in Adj-inconsistent units' pictures, *p*s $< .001$. The bar plots of the fixation proportions with the classifier *gen* and *kuai* are illustrated in Figure 4-4B and 4C respectively.

5.2. Behavioural data

In analysing participants' behavioural data, their choices of the target picture among the four pictures in each trial in different conditions with CCNs/CMNs were calculated. Participants' behavioural data were analysed in the same way as their fixation proportions. Participants' choices for the final target picture are summarized in Table 3, with results of the statistical analysis shown in the Appendix.

In general, participants' behavioural data reveal different patterns for the classifier *ba* and the classifiers *gen* and *kuai*. With the baseline structure [Cl-N], there was no important difference: for all three classifiers, most participants (>80%) chose the 'objects measured in individual units' picture as the target picture when the noun was a CCN, while the two CMN-denoting pictures were each selected half of the time when the noun was a CMN. With the neutral structure [Cl-A-N], most participants (>80%) chose the 'objects measured in individual units' picture when the noun was a CCN regardless of classifiers. When the noun was a CMN, most of them (>60%) chose the 'entities measured in Adj-inconsistent units' picture with *ba*, but chose the 'entities measured in Adj-consistent units' picture with *gen* and *kuai*. With the critical structure [A-Cl-N], the classifier made no major difference when the noun was a CMN: most participants chose the 'entities measured in Adj-consistent units' picture (>60%). But when the noun was a CCN, 95% of participants chose the 'objects measured in groups' picture with *ba*, while 95% of them chose the 'objects measured in individual units' picture with *gen* and *kuai*.

Table 3

Behavioural results (Since participants' responses were similar for gen and kuai, their final choices with these two classifiers were collapsed.)

	Baseline [Cl-N]		**Neutral [Cl-A-N]**		**Critical [A-Cl-N]**	
	CCN	**CMN**	**CCN**	**CMN**	**CCN**	**CMN**
Ba						
Objects measured in individual units	**84%**	0%	**80%**	0%	5%	0%
Objects measured in groups	16%	0%	20%	0%	**95%**	0%
Entities measured in Adj-inconsistent units	0%	45%	0%	**60%**	0%	20%
Entities measured in Adj-consistent units	0%	**55%**	0%	40%	0%	**80%**
Gen/Kuai						
Objects measured in individual units	**95%**	0%	**95%**	0%	**95%**	0%
Objects measured in portions	5%	0%	5%	0%	5%	0%
Entities measured in Adj-inconsistent units	0%	50%	0%	20%	0%	40%
Entities measured in Adj-consistent units	0%	50%	0%	**80%**	0%	**60%**

To sum up, the manipulation of nominal phrase structures evoked L2ers' different fixation patterns. Especially with the critical structure [A-Cl-N], on hearing a classifier right after an adjective, participants landed more fixations on the mass/plural-denoting pictures than the singular/individual-denoting picture. This pattern indicates that they used the [A-Cl] structure to interpret the classifier to be a massifier, and build a mass/plural anticipation for the upcoming noun. Furthermore, participants' fixation patterns in the ROI of nouns are highly consistent with their behaviour, which confirms the validity of the data.

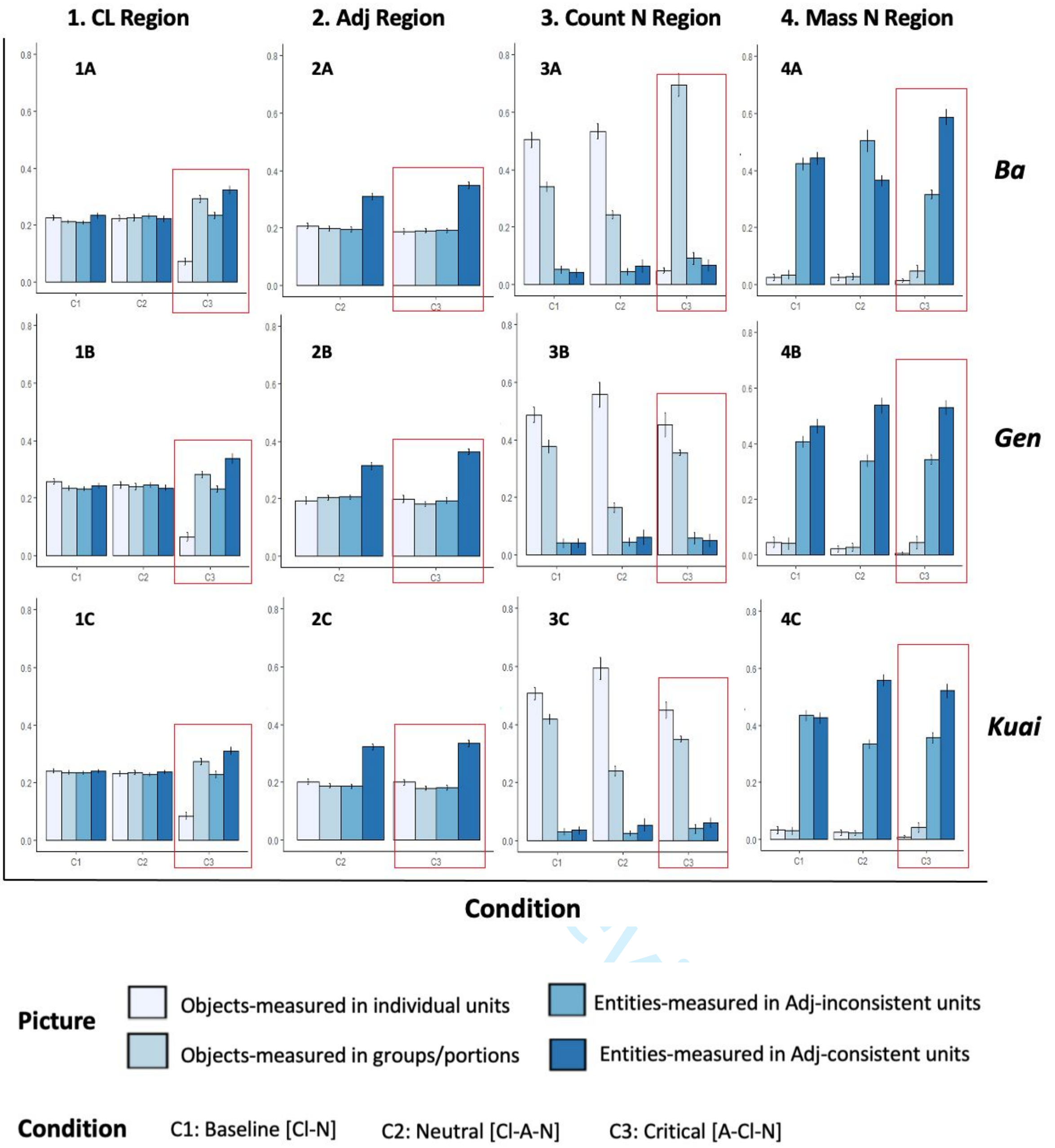

Figure 4
*Fixation proportions in the ROIs (1-Classifier Region, 2-Adjective Region, 3-CCN Region, and 4-CMN Region). When the classifier is **ba**, the corresponding plots are 1A, 2A, 3A and 4A. When the classifier is **gen**, the corresponding plots are 1B, 2B, 3B and 4B. When the classifier is **kuai**, the corresponding plots are 1C, 2C, 3C and 4C. In each plot, the lateral axis represents conditions of nominal phrases, the vertical axis represents fixation proportions. Bars in different colours represent fixation proportions in different pictures. Condition 3 in each plot was highlighted with a rectangle since it is the one with the massifier structure (critical condition).*

5.3. L1ers vs. L2ers

To further explore whether L2 participants exhibit native-like patterns in using massifier-biased syntactic cues during on-line processing, participants' data in the current study were compared to native speakers' data in Yao et al. (2022). We particularly focused on the distributions of fixations among the four pictures before the output of the nouns. These data reveal how participants incrementally interpret the nominal phrases based on each item they heard. We compared L2ers' patterns with L1ers' in each condition by collapsing across different classifiers and CCNs/CMNs. Following Mirman et al. (2008), we used Growth Curve plots to illustrate participants' fixation distributions in real time processing. The proportions of fixations on each picture along the time sequence of the nominal phrases were calculated in each condition. A sample was taken every 50 milliseconds (ms). In each audio nominal phrase, the averaged duration was 300ms for the numeral, 250ms for the classifier, 300ms for the adjective, and 350ms for the noun. We used Linear Regression Modeling in the ROI of classifier and adjective in each condition to test whether the fixation distribution patterns differ significantly between L1ers and L2ers. The model included simulated growth curve (generated with *poly*), picture (the four types of visual display) and participant groups (native speakers and L2ers) as fixed effects. The interaction between growth curve and participant groups would indicate whether the fixation distribution patterns differ significantly between native speakers and L2ers. All the code and data are available at *osf.io/5fzxg*.

5.3.1. With the baseline structure [Cl-N]

The proportions of fixations on the four pictures for L1ers and L2ers are plotted in Figure 5, where the top plot presents L1ers' fixation distribution patterns, while the bottom one presents L2ers'. With the structure [Cl-N], there is only one ROI before the noun: the classifier region, which is represented by the blue rectangle. Rectangles with dashed lines represent the duration (from the onset to the offset) of each item: the numeral, the classifiers and the nouns.

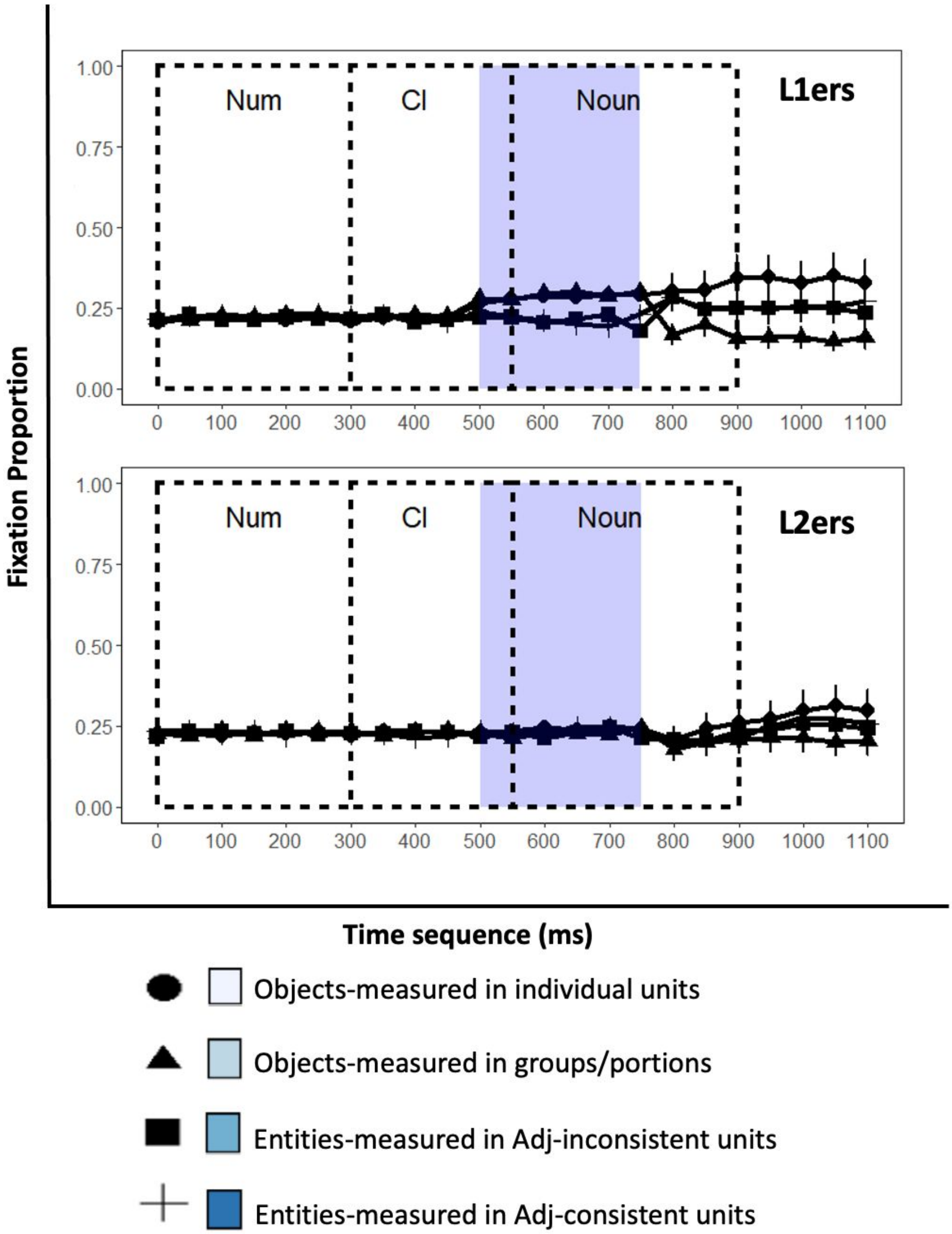

Figure 5
Fixation distributions of L1ers and L2ers hearing the baseline structure [Cl-N]

For both L1ers and L2ers, the proportions of fixations on the four pictures stayed around the random level (0.25) before they reached the nouns. The interaction between growth curves and participant groups was not significant, $p = .075$. This observation is reasonable considering the structure [Cl-N] is a baseline structure which contains no mass/count syntactic cue that can be used to build corresponding anticipations. Both L1ers and L2ers could only decide which is the target picture once they heard the nouns.

5.3.2. With the neutral structure [Cl-A-N]

The proportions of fixations on the four pictures for L1ers and L2ers are plotted in Figure 6. With the structure [Cl-A-N], there are two ROIs before the noun: the classifier region and the adjective region, which are represented by the blue rectangle and the red rectangle respectively.

In the ROI of the classifier, both L1ers and L2ers randomly fixated on the four pictures. The interaction between growth curves and participant groups was not significant, $p = .878$. In the ROI of the adjective, however, L1ers and L2ers exhibited different patterns: the interaction between growth curves and participant groups was significant, $p <.001$. The onset of the adjectives did not affect L1ers' fixation distributions: the fixation proportions on the four pictures stayed around the random level in the adjective region. On the other hand, on hearing the adjective, L2ers started to land more fixations on the 'entities measured in Adj-consistent units' picture, in which the size of the group/unit is either the biggest or the smallest among the four pictures, and always consistent with the adjective.

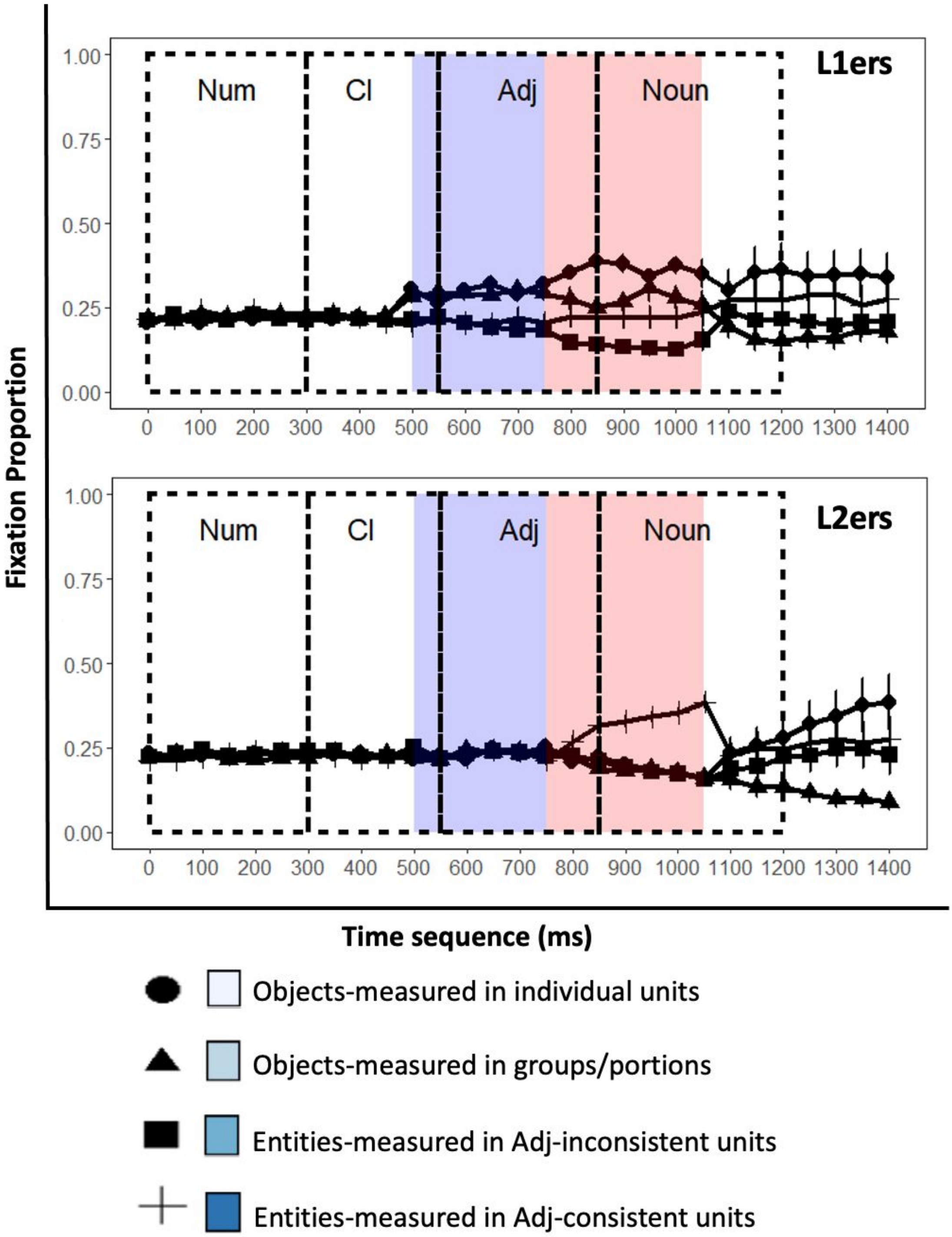

Figure 6

Fixation distributions of L1ers and L2ers hearing the neutral structure [Cl-A-N]

5.3.3. With the critical structure [A-Cl-N]

The proportions of fixations on the four pictures for L1ers and L2ers are plotted in Figure 7. With the structure [A-Cl-N], there are two ROIs before the noun: the adjective region and the classifier region, which are represented by the red rectangle and the blue rectangle respectively.

In the adjective region, L1ers and L2ers behaved differently: the interaction between growth curves and participant groups was significant, $p < .005$. For L1ers, the onset of the adjective drove their attention to the mass/plural-expressing pictures. For L2ers, however, similar to their patterns in the neutral condition, on hearing the adjective, they directed their fixations to the 'entities measured in Adj-consistent units' picture. In the classifier region, L1ers and L2ers also behaved differently: the interaction between growth curves and participant groups was significant, $p < .001$. Even though they both exhibited the preference for the mass/plural-expressing pictures, L2ers were slower than L1ers. On hearing the adjective directly following the numeral, L1ers immediately predicted that the upcoming item would be a massifier, and consequently anticipated the nominal to have a mass/plural interpretation. As a result, more fixations were landed on the mass/plural-expressing pictures immediately following the onset of the adjective. L2ers did not exhibit this preference for the mass/plural-expressing pictures on hearing adjectives. The adjective directly following the numeral did not trigger their expectations for a massifier or a mass/plural interpretation. Instead, it drove their attention to the picture in which the objects are organized in an adjective-consistent size. Only on hearing the classifier following the adjective, did L2ers start to shift their attention to the mass/plural-expressing pictures, consistent with the massifier-biased syntactic cue, although at a later point than L1ers.

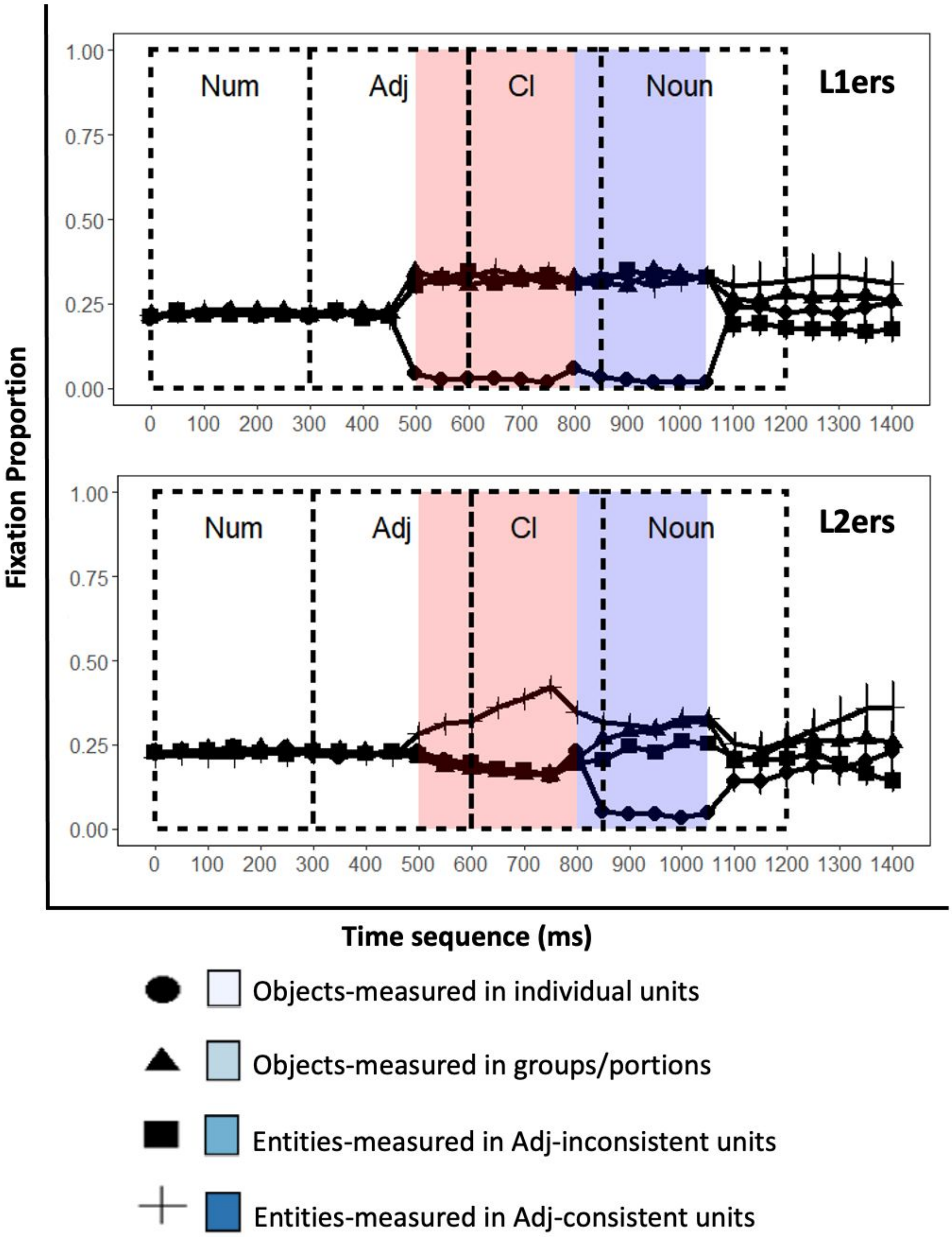

Figure 7.

Fixation distributions of L1ers and L2ers hearing the critical structure [A-Cl-N]

Discussion

In general, there are two main findings for L2 participants in the current study: (1) they can automatically use the massifier-biased syntactic cue [Num-Adj-Cl] to make a mass/plural-preferred interpretation in real time processing; (2) they were slower than L1ers in the predictive use of this massifier-biased syntactic cue. We discuss these two main findings one by one in the following part.

1. L2ers' use of the massifier-biased syntactic cue

The fixation proportions among the four pictures in the classifier region revealed that participants in the current study can automatically use the structure [Num-Adj-Cl] as a massifier-biased syntactic cue to make mass/plural-preferred interpretations in real time processing. To be specific, with the baseline and neutral structures [Cl-N] and [Cl-A-N], on hearing the classifier, participants randomly fixated on the four pictures without any specific preference. However, on hearing the critical structure [A-Cl], participants landed more fixations on the mass/plural-expressing pictures than the 'objects measured in individual units' picture. The preference for mass/plural-expressing pictures on hearing the classifiers in the critical condition compared to the baseline and neutral conditions indicated that participants interpreted the classifier in the structure [A-Cl] as a massifier, which consequently requires a mass/plural-denoting nominal. Recall that the classifiers used in the current study are all ambiguous between count-classifier and massifier uses, so hearing the actual classifier was not in itself a cue to interpretation. Participants' different fixation patterns between the critical and baseline/neutral conditions indicated that they were aware of the two different meanings (massifier and count-classifier) of these dual-role classifiers, and could take advantage of the massifier-biased syntactic cue to select the appropriate massifier meaning of these dual-role classifiers. This is similar to L1ers' fixation patterns in the classifier region.

L2ers' native-like patterns can also be found in the fixation proportions in the noun region. When hearing a CCN, different classifiers yield different patterns with the critical structure [A-Cl-N]. When the classifier was *ba*, participants directed significantly more fixations to the 'objects measured in groups' picture. This is consistent with the observations with L1ers that the classifier *ba* in the structure [A-Cl] gives rise to a nominal-with-a-plural-set meaning. With the classifiers *gen* and *kuai*, which are dividers when they are interpreted as massifiers, participants overwhelmingly switched their fixations away from the mass-nominal-denoting pictures towards the 'objects measured in individual units' picture, which preserves the 'whole/indivisible object' interpretation of the CCN, but is inconsistent with the syntax of the phrase they were listening to. These patterns with *gen* and *kuai* were also observed with L1ers, and indicated that the massifier-biased syntactic structure is not enough to coerce a mass interpretation for CCNs, especially when there was a conflict between semantic properties of classifiers and CCNs. To be specific, when taking the role of a massifier, *gen* and *kuai* create a certain measurement by dividing the entities denoted by the nouns they precede, but the entities denoted by the CCNs cannot be divided, creating a conflict. Participants' target picture selection (the behavioural data) is consistent with their fixation distributions in the noun region, which is also consistent with L1ers' behavioural patterns. These native-like patterns of L2ers indicate that the dual-role classifiers and massifier-biased syntactic cues which are unique-to-L2 constructions have been implicitly acquired by high-proficiency L2ers, and can be automatically used to build anticipations for the upcoming nouns.

It should be noted that L2ers' processing of nominal phrases may also be affected by their understandings of the classifiers, which should be consistent with these classifiers' co-occurrence frequency with CMNs vs. CCNs in daily life. Using two Mandarin Chinese corpora (National Language Resources Monitoring and Research Centre, Broadcast Media Language Branch, 2009; Ministry of Education and institute of Applied Linguistics, 2009) and the number of Baidu (Mandarin Chinese version of Google) hits (May 2019) (Pollatsek et al., 2010), the co-occurrence frequency counts of the CCNs and CMNs used in the current research with the three classifiers were calculated, and are illustrated in Table 4.

Table 4.
The mean co-occurrence frequency (per million) of the three classifiers

	Ba			***Gen***			***Kuai***		
	CCN	CMN	*p*	CCN	CMN	*p*	CCN	CMN	*p*
Frequency	0.68	0.64	.35	0.35	0.51	<.01	0.43	0.59	<.01

From Table 4 we can see that the classifier *gen* and *kuai* are much more likely to occur with CMNs than CCNs, while the classifier *ba* was equally likely to occur with either CCNs or CMNs. If participants interpreted nominals by using their understandings of the dominant meanings of the classifiers instead of the mass/count syntactic cues, in all three conditions on hearing the classifier *gen* or *kuai*, they should have fixated on the CMN-denoting pictures but not the CCN-denoting pictures, which is not what we observed. The clear and obvious preference for the mass/plural-expressing pictures with the critical structure [A-Cl-N] on hearing the classifiers indicates that Dutch-Mandarin learners have acquired the dual-role classifiers and the massifier-biased syntactic cue, and can predictively use this knowledge in real time processing.

2. L2ers' slowness of using the massifier-biased syntactic cue

When directly comparing L2ers' data to L1ers' we found that L2ers were slower than L1ers in making use of the [A-Cl] cue to build mass/plural predictions: on hearing the adjective directly after the numeral, L1ers immediately predicted the upcoming item to be a massifier and landed their fixations on the mass/plural-expressing pictures, while L2ers started to direct their attention to these pictures only after the onset of the classifier. In a sense, L1ers were sensitive to the [Num-Adj] word order, and used it as a predictor of an upcoming massifier, whereas L2ers seemed to need to actually hear the classifier to generate the massifier structure, even though the classifier itself was not informative to either a mass/plural or count interpretation.

We speculate that the slowness of L2ers in making use of the [A-Cl-N] structure to build mass/plural-preferred anticipations is not just the general processing slowness of L2ers found in a number of previous studies (Hahne & Friederici, 2001; Lew-Williams & Fernald, 2010; Sanders & Neville, 2003). Rather, we argue that the later fixation shift to mass/plural-expressing pictures with the critical structure [A-Cl-N] exhibited by L2ers compared to L1ers reveals L2ers' different processing strategies for adjectives. Compared to L1ers, who were sensitive to the syntactic position of the adjective, L2ers were more sensitive to the semantic information encoded in it. In both the neutral condition and the critical condition, on hearing adjectives, L2ers shifted fixations to the 'entities measured in Adj-consistent units' picture in which the size of the group/unit is either the biggest or the smallest among the four pictures, and always consistent with the adjective. This pattern indicates that L2ers automatically used the semantic information of adjectives to interpret nominal phrases and direct their fixations. L1ers, on the other hand, either waited for the onset of nouns to react (in the neutral condition) or used the occurrence of adjectives directly following the numeral initially as a syntactic cue for massifiers (in the critical condition), only integrating the adjective's semantic contribution in the noun region. The different processing patterns of adjectives for L2ers compared to L1ers are consistent with previous studies which found that late L2ers are sensitive to semantic information over syntactic information (Felser et al., 2003; Lau & Grüter, 2015; Lew-Williams & Fernald, 2009; Papadopoulou & Clahsen, 2003; Roberts & Felser, 2011). Clahsen & Felser (2018) claimed that lexical, semantic and pragmatic (or other types of nongrammatical information) may take a priority in L2 processing, since L2ers potentially are more sensitive to these types of information than L1ers. Gr ü ter et al. (2020) found a greater reliance on semantic cues compared to syntactic cues in L2 processing. Based on these existing findings, it is reasonable to observe L2ers' prioritizing the semantic information over the syntactic information supplied by adjectives in real time processing.

However, it should be noted that in the current study, even though L2ers were more sensitive to the semantic meanings of adjectives than their interpretive syntactic information, they did quickly 'catch up' and fixated on the mass/plural-expressing pictures in the critical condition upon hearing the classifiers. This suggests that L2 participants were sensitive to the massifier-biased syntactic cue and the corresponding mass/plural interpretation, but were distracted by the semantic content of the adjective and needed more time, and, possibly, the confirmation of the whole [Num-Adj-Cl] sequence, to make a mass/plural prediction.

To conclude, the current study conducted a VWP experiment to explore whether high-proficiency late Dutch-Mandarin learners can predictively use dual-role classifiers and a massifier-biased syntactic cue which are unique-to-L2 constructions and can only be acquired implicitly. The results revealed that although Dutch-Mandarin learners are more sensitive to the semantic information of adjectives compared to native Mandarin speakers, they can predictively use the [A-Cl-N] structure to make a mass/plural-preferred interpretation in real time processing. This native-like pattern indicates that they have acquired these unique-to-L2 constructions through implicit learning and can automatically use them to build anticipations.

Declarations

Conflicts of interest/competing interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Funding

This work was supported by CSC (China Scholarship Council), the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 613465: AThEME (Advancing the European Multilingual Experience), Beijing Social Science Funds (22YYC016) and Science Foundation of Beijing Language and Culture University (supported by "the Fundamental Research Funds for the Central Universities") (22YJ250001) to the first author.

References

Allan, K. (1977). Classifiers. *Language*, 53, 285-311.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38: 419−439.

Altmann, G., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive science*, 33(4), 583-609.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.

Barner, D., Inagaki, S., & Li, P. (2009). Language, thought, and real nouns. *Cognition*, 111(3), 329-344.

Barr, D. J., Gann, T. M., and Pierce, R. S. (2011). Anticipatory baseline effects and information integration in visual world studies. *Acta Psychologica*, 137: 201–07.

Borer, H. (2005). *In name only* (Vol. 1). Oxford University Press on Demand.

Cheng, C. Y. (1973). Response to Moravscik. In Hintikka, J., Moravcsik, J., & Suppes, P. (eds), *Approaches to natural language*. 286-288. Dordrecht: Reidel.

Cheng, L. L. S., & Sybesma, R. (1998). Yi-wan tang, yi-ge tang: Classifiers and massifiers. *Tsing Hua journal of Chinese studies*, 28(3), 385-412.

Cheng, L. L. S., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic inquiry*, 30(4), 509-542.

Cheng, L. L. S., & Sybesma, R. (2012). Classifiers and DP. *Linguistic inquiry*, 43(4), 634-650.

Cheng, L. L. S., Doetjes, J., & Sybesma, R. (2008). How universal is the Universal Grinder. *Linguistics in the Netherlands*, 2008, 50-62.

Chierchia, G. (1998a). *Plurality of mass nouns and the notion of "semantic parameter"* (pp. 53-103). Springer Netherlands.

Chierchia, G. (1998b). Reference to kinds across language. *Natural language semantics*, *6*(4), 339-405.

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese*, 174(1), 99-149.

Choi, S.H., Ionin, T., and Zhu, Y. Q. (2017). L1 Korean and L1 Mandarin L2 English learners' acquisition of the count/mass distinction in English. *Second language research*, *34*(2), 1-31.

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*(3), 693-706.

Cleary, A. M., & Langley, M. M. (2007). Retention of the structure underlying sentences. *Language and Cognitive Processes*, 22, 614–628.

Doetjes, J. (1997). Quantifiers and selection. On the distribution of quantifying expressions in French, Dutch and English, The Hague: Holland Academic Graphics (HIL dissertation series).

Donaldson, B. (2011). Left dislocation in near-native French. *Studies in Second Language Acquisition, 33*(3), 399-432.

Engelhardt, P. E. & Ferreira, F. (2010). Processing coordination ambiguity. *Language and speech*, 53(4), 494-509.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, 24(3), 453.

Gleitman, L., & Papafragou, A. (2005). Language and Thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning*. 633-661. Cambridge University Press.

Gong, J. S. (2010). Chinese classifier acquisition: comparison of L1 child and L2 adult development. *Theses, Dissertations, Professional Papers*: 192.

Grüter, T., & Rohde, H. (2013). L2 processing is affected by RAGE: Evidence from reference resolution. *In the 12th conference on Generative Approaches to Second Language Acquisition* (GASLA).

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, *35*(2), 221-234.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191-215.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4, 123–141.

Hopp, H. (2012). The on-line integration of inflection in L2 processing: Predictive processing of German gender. In A. K. Biller, E. Y. Chung & A. E. Kimball (eds.), *Proceedings of the 36th Boston University Conference on Language Development* (BUCLD 36), pp. 226–245. Somerville, MA: Cascadilla Press.

Hopp, H. (2015). Semantics and morphosyntax in L2 predictive sentence processing. *International Review of Applied Linguistics*, 53, 277–306.

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, 32(2), 277-307.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137: 151-171.

Iljic, R. (1994). Quantification in Mandarin Chinese: two markers of plurality. *Linguistics*, 32(1), 91-116.

Kamide Y (2008) Anticipatory processes in sentence processing. *Language and Linguistics Compass,* (2), 647–70.

Klein, N. M., Carlson, G. N., Li, R. J., Jaeger, T. F., & Tanenhaus, M. K. (2012). Classifying and massifying incrementally in Chinese language comprehension. *Count and mass across languages*, 261-282.

Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. *Lexical Matters*, (eds.) Ivan A. Sag and Anna Szablocsi. Stanford, CA: Center for the Study of Language and Information, 29-53.

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, 119-142.

Lau, E., & Grüter, T. (2015). Real-time Processing of Classifier Information by L2 Speakers of Chinese. *BUCLD 39 Proceedings*, Cascadilla Press.

Leung, J. H., & Williams, J. N. (2014). Cross-linguistic differences in implicit language learning. *Studies in second language acquisition*, 36(04), 733-755.

Lew-Williams, C., & Fernald, A. (2009). Fluency in using morphosyntactic cues to establish reference: How do native and non-native speakers differ. *In Proceedings of the 33rd annual Boston university conference on language development*, Vol. 1, 290-301.

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, 63(4), 447-464.

Li, P., Barner, D., & Huang, B. (2008). Classifiers as count syntax: Individuation and measurement in the acquisition of Mandarin Chinese. *Language Learning and Development*, 4(4), 1-42.

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology, 58*, 487-524.

Liang, S. Y. (2008). The acquisition of Chinese nominal classifiers by L2 adult learners. In Majorie K. M. Chan & Hana Kang (Eds.), *Proceedings of the 20th North American Conference on Chinese Linguistics* (NACCL-20) (pp. 309-326). Columbus, OH: Ohio State University.

Mirman, D., Dixon, J. A., and Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: growth curves and individual differences. *Journal of Memory and Language*, *59*, 475–494.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers. *Studies in second language acquisition*, *25*(3), 351-398.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing. *Studies in Second Language Acquisition*, *25*(4), 501-528

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329-347.

Polio, C. (1994). Non-native speakers' use of nominal classifiers in Mandarin Chinese. *Journal of the Chinese Language Teachers Association*, *29*(3), 51-66.

Pollatsek, A., Drieghe, D., Stockall, L., & de Almeida, R. G. (2010). The interpretation of ambiguous trimorphemic words in sentence context. *Psychonomic Bulletin & Review*, 17(1), 88-94.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 317–327.

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, *33*(4), 829-856.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguist*, *32*(2), 299-331.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, & Hernstadt (1991) and Knowlton & Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, *27*, 235–268.

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing II: Segmentation, semantics, and syntax in non-native speakers. *Cognitive Brain Research*, *15*, 215–227.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*,110, 104038.

Slabakova, R. (2006). Is there a critical period for semantics? *Second Language Research*, *22* (3), 302-338.

Tai, J. H. (1992). Variation in classifier systems across Chinese dialects: towards a cognition-based semantic approach. *Chinese Language and Linguistics*, *1*, 587-608.

Tai, J. H. (1994). Chinese classifier systems and human categorization. In *In hornor of William S. Y. Wang: Interdisciplinary studies on language and language change*. Ed. M.Y. Chen and O. Tzeng, 479-494. Taipei: Pyramid Press.

Tai, J., & Wang, L. (1990). A Semantic Study of the Classifier Tiao. *Journal of the Chinese Language Teachers Association*, *25*(1), 35-56.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning*, *64*(2), 279-309.

Trenkic, D., Mirkovic, J., & Altmann, G. T. M. (2014). Real-time grammar processing by native and non-native speakers: constructions unique to the second language. *Bilingualism: Language and Cognition*, *17*(2), 237-257.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer Science & Business Media

Yan, H., Liu, L., Tate, M., Wang, L., Bin, Y., & Zhu, X. (2009). Edexcel GCSE Chinese. United Kingdom: Pearson Education Limited

Yao, P., Stockall, L., Hall, D., & Borer, H. (2022). Processing evidence for the grammatical encoding of the mass/count distinction in Mandarin Chinese. *Journal of Psycholinguistic Research, 51*(2), 341-371.

Yip, M. (2002). *Tone*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press. *DOI.10.1017/CBO9781139164559*

Zhang, N. N. (2012). Countability and numeral classifiers in Mandarin Chinese. A Cross Linguistic Exploration of the Count Mass Distinction, 220-237.

Appendix

Noun Rating Test

A Noun Rating Test was conducted to select count and mass nouns based on native Mandarin speakers' judgments. Among a word pool containing 90 high frequency two-character nouns which can normally occur with one of the three classifiers (selected from National Language Resources Monitoring and Research Centre, Broadcast Media Language Branch, 2009; Ministry of Education and institute of Applied Linguistics, 2009), ten native Mandarin speakers were asked to rate each noun for 'divisibility' (Cheng, 1973; Krifka 1992), using a 5-point scale in which 1 stands for 'divisible' while 5 stands for 'indivisible'. They were told that a noun is 'divisible' if the entity denoted by it can be divided several times, and each part of it after being divided still has the original property. On the contrary, a noun is 'indivisible' if the entity denoted by it cannot be divided, or each part after being divided possesses different features from the original entity. Krifka (1992) argued that entities which are divisible are mass and entities which are not divisible are count. Based on the rating results, 12 CMNs (mean rating score = 1.58) and 12 CCNs (mean rating score = 4.39) were chosen. These words are summarized in Table A.

Table A.
The count & mass nouns selected in the Noun Rating Test

	Count nouns		Mass nouns	
Ba	*Shaozi*	'spoon'	*Shizi*	'pebble'
	Chizi	'ruler'	*Zhongzi*	'seed'
	Tongsuo	'locker'	*Hongdou*	'red bean'
	Yaoshi	'key'	*Muchai*	'firewood'
Gen	*Huanggua*	'cucumber'	*Shengzi*	'string'
	Xiangjiao	'banana'	*Xiangchang*	'sausage'
	Muahua	'fried twist'	*Mugun*	'stick'
Kuai	*Yupei*	'jade'	*Dangao*	'cake'
	Jimu	'building block'	*Huangyou*	'butter'
	Shoujuan	'handkerchief'	*Nailao*	'cheese'
	Huaban	'sketchpad'	*Xiangpi*	'eraser'
	Hongzhuan	'brick'	*Feizao*	'soap'

Table B.

Linear regression for the fixation proportions in the ROI of classifier when classifier is ba

Contrast	Estimate	SE	t	*p*
(Intercept)	0.22	0.00	77.73	<.001
[Cl-N]				
‘Objects in individual units’ - ‘Objects in groups’	0.01	0.01	0.93	.36
‘Objects in groups’ - ‘Entities in Adj-inconsistent units’	0.00	0.01	0.16	.87
‘Entities in Adj-inconsistent units’- ‘Entities in Adj-consistent units’	-0.02	0.01	-1.60	.11
[Cl-A-N]				
‘Objects in individual units’ - ‘Objects in groups’	0.00	0.01	-0.16	.87
‘Objects in groups’ - ‘Entities in Adj-inconsistent units’	-0.01	0.01	-0.38	.70
‘Entities in Adj-inconsistent units’- ‘Entities in Adj-consistent units’	0.01	0.01	0.69	.49
[A-Cl-N]				
‘Objects in individual units’ - ‘Objects in groups’	-0.22	0.01	-15.59	<.001
‘Objects in groups’ - ‘Entities in Adj-inconsistent units’	0.06	0.01	4.04	<.001
‘Entities in Adj-inconsistent units’- ‘Entities in Adj-consistent units’	-0.09	0.01	-6.25	<.001

Table C.

Linear regression for the fixation proportions in the ROI of classifier when classifier is gen

Contrast	Estimate	SE	t	*p*
(Intercept)	0.24	0.00	73.84	<.001
[Cl-N]				
‘Objects in individual units’ - ‘Objects in portions’	0.02	0.02	1.47	.14
‘Objects in portions’ - ‘Entities in Adj-inconsistent units’	0.00	0.02	0.18	.86
‘Entities in Adj-inconsistent units’- ‘Entities in Adj-consistent units’	-0.01	0.02	-0.72	.47
[Cl-A-N]				
‘Objects in individual units’ - ‘Objects in portions’	0.01	0.02	0.35	.72
‘Objects in portions’ - ‘Entities in Adj-inconsistent units’	-0.01	0.02	-0.37	.71
‘Entities in Adj-inconsistent units’- ‘Entities in Adj-consistent units’	0.01	0.02	0.69	.49
[A-Cl-N]				
‘Objects in individual units’ - ‘Objects in portions’	-0.22	0.02	-13.77	<.001
‘Objects in portions’ - ‘Entities in Adj-inconsistent units’	0.05	0.02	3.20	<.001

'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.11	0.02	-6.75	<.001

Table D.

Linear regression for the fixation proportions in the ROI of classifier when classifier is kuai

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.00	90.98	<.001
[Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.01	0.01	0.47	.64
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.00	0.01	0.11	.91
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.01	0.01	-0.56	.57
[Cl-A-N]				
'Objects in individual units' - 'Objects in portions'	0.00	0.01	-0.29	.77
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.01	0.01	0.55	.58
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.01	0.01	-0.70	.49
[A-Cl-N]				
'Objects in individual units' - 'Objects in portions'	-0.19	0.01	-15.24	<.001
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.05	0.01	3.68	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.08	0.01	-6.60	<.001

Table E.

Linear regression for the fixation proportions in the ROI of adjective when classifier is ba

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.00	69.42	<.001
[Cl-A-N]				
'Objects in individual units' - 'Objects in groups'	0.01	0.01	0.69	.49
'Objects in groups' - 'Entities in Adj-inconsistent units'	0.00	0.01	0.15	.88
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.11	0.01	-8.54	<.001
[A-Cl-N]				
'Objects in individual units' - 'Objects in groups'	0.00	0.01	-0.15	.88
'Objects in groups' - 'Entities in Adj-inconsistent units'	0.00	0.01	-0.08	.94
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.16	0.01	-12.07	<.001

Table F.

Linear regression for the fixation proportions in the ROI of adjective when classifier is gen

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.00	63.07	<.001
[Cl-A-N]				
'Objects in individual units' - 'Objects in portions'	-0.01	0.01	-0.70	.49
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.00	0.01	-0.16	.87
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.11	0.01	-7.45	<.001
[A-Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.02	0.01	1.20	.23
'Objects in portions' - 'Entities in Adj-inconsistent units'	-0.01	0.01	-0.83	.41
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.17	0.01	-11.65	<.001

Table G.

Linear regression for the fixation proportions in the ROI of adjective when classifier is kuai

Contrast	Estimate	SE	t	*p*
(Intercept)	0.22	0.00	69.18	<.001
[Cl-A-N]				
'Objects in individual units' - 'Objects in portions'	0.01	0.01	1.07	.29
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.00	0.01	0.12	.90
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.14	0.01	-10.49	<.001
[A-Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.02	0.01	1.59	.11
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.00	0.01	-0.27	.79
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.15	0.01	-11.88	<.001

Table H.

Linear regression for the fixation proportions in the ROI of count noun when classifier is ba

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.01	37.62	<.001
[Cl-N]				
'Objects in individual units' - 'Objects in groups'	0.16	0.03	5.56	<.001
'Objects in groups' - 'Entities in Adj-inconsistent units'	0.29	0.03	9.79	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	0.01	0.03	0.39	.70
[Cl-A-N]				

'Objects in individual units' - 'Objects in groups'	0.29	0.03	9.87	<.001
'Objects in groups' - 'Entities in Adj-inconsistent units'	0.20	0.03	6.78	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.02	0.03	-0.70	.48
[A-Cl-N]				
'Objects in individual units' - 'Objects in groups'	-0.65	0.03	-21.96	<.001
'Objects in groups' - 'Entities in Adj-inconsistent units'	0.60	0.03	20.51	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	0.02	0.03	0.80	.43

Table I.

Linear regression for the fixation proportions in the ROI of count noun when classifier is gen

Contrast	Estimate	SE	t	*p*
(Intercept)	0.22	0.01	30.97	<.001
[Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.11	0.04	3.05	<.001
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.34	0.04	9.54	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	0.00	0.04	0.00	1.00
[Cl-A-N]				
'Objects in individual units' - 'Objects in portions'	0.39	0.04	11.08	<.001
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.12	0.04	3.40	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.02	0.04	-0.50	.62
[A-Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.10	0.04	2.75	.01
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.30	0.04	8.40	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	0.01	0.04	0.26	.80

Table J.

Linear regression for the fixation proportions in the ROI of count noun when classifier is kuai

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.01	40.85	<.001
[Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.09	0.03	3.14	<.001
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.39	0.03	13.91	<.001

'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.01	0.03	-0.18	.86
[Cl-A-N]				
'Objects in individual units' - 'Objects in portions'	0.35	0.03	12.62	<.001
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.22	0.03	7.72	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.03	0.03	-1.03	.30
[A-Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.10	0.03	3.58	<.001
'Objects in portions' - 'Entities in Adj-inconsistent units'	0.31	0.03	11.07	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.02	0.03	-0.74	.46

Table K.

Linear regression for the fixation proportions in the ROI of mass noun when classifier is ba

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.01	41.54	<.001
[Cl-N]				
'Objects in individual units' - 'Objects in groups'	-0.01	0.03	-0.36	.72
'Objects in groups' - 'Entities in Adj-inconsistent units'	-0.39	0.03	-14.16	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.02	0.03	-0.68	.50
[Cl-A-N]				
'Objects in individual units' - 'Objects in groups'	0.00	0.03	-0.11	.91
'Objects in groups' - 'Entities in Adj-inconsistent units'	-0.48	0.03	-17.33	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	0.14	0.03	5.06	<.001
[A-Cl-N]				
'Objects in individual units' - 'Objects in groups'	-0.03	0.03	-1.19	.23
'Objects in groups' - 'Entities in Adj-inconsistent units'	-0.27	0.03	-9.77	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.27	0.03	-9.86	<.001

Table L.

Linear regression for the fixation proportions in the ROI of mass noun when classifier is gen

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.01	41.23	<.001
[Cl-N]				

'Objects in individual units' - 'Objects in portions'	0.00	0.03	0.17	.87
'Objects in portions' - 'Entities in Adj-inconsistent units'	-0.37	0.03	-13.25	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.06	0.03	-2.03	.04
[Cl-A-N]				
'Objects in individual units' - 'Objects in portions'	0.00	0.03	-0.17	.87
'Objects in portions' - 'Entities in Adj-inconsistent units'	-0.31	0.03	-11.24	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.20	0.03	-7.17	<.001
[A-Cl-N]				
'Objects in individual units' - 'Objects in portions'	-0.04	0.03	-1.43	.15
'Objects in portions' - 'Entities in Adj-inconsistent units'	-0.30	0.03	-10.73	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.19	0.03	-6.81	<.001

Table M.

Linear regression for the fixation proportions in the ROI of mass noun when classifier is kuai

Contrast	Estimate	SE	t	*p*
(Intercept)	0.23	0.00	50.82	<.001
[Cl-N]				
'Objects in individual units' - 'Objects in portions'	0.00	0.02	0.12	.90
'Objects in portions' - 'Entities in Adj-inconsistent units'	-0.41	0.02	-18.11	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	0.01	0.02	0.35	.73
[Cl-A-N]				
'Objects in individual units' - 'Objects in portions'	0.00	0.02	0.12	.90
'Objects in portions' - 'Entities in Adj-inconsistent units'	-0.31	0.02	-13.95	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.22	0.02	-10.05	<.001
[A-Cl-N]				
'Objects in individual units' - 'Objects in portions'	-0.03	0.02	-1.45	.15
'Objects in portions' - 'Entities in Adj-inconsistent units'	-0.32	0.02	-14.13	<.001
'Entities in Adj-inconsistent units'- 'Entities in Adj-consistent units'	-0.16	0.02	-7.35	<.001

Table N

Linear regression for behaviour data when classifier is ba

	Estimate	SE	t	*p*
Count nouns				
(Intercept)	0.25	0.012	21.068	<.001
[Cl-N]				
objects measured in individual units - objects measured in groups	0.667	0.057	11.691	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.167	0.057	2.923	<.005
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.057	0	1
[Cl-A-N]				
objects measured in individual units - objects measured in groups	0.647	0.059	11.027	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.176	0.059	3.007	<.005
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.059	0	1
[A-Cl-N]				
objects measured in individual units - objects measured in groups	-0.882	0.059	-15.037	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.941	0.059	16.04	<.001
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.059	0	1
Mass nouns				
(Intercept)	0.25	0.016	15.49	<.001
[Cl-N]				
objects measured in individual units - objects measured in groups	0	0.08	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.471	0.08	-5.896	<.001
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	-0.059	0.08	-0.737	0.462
[Cl-A-N]				
objects measured in individual units - objects measured in groups	0	0.08	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.647	0.08	-8.107	<.001
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0.294	0.08	3.685	<.001
[A-Cl-N]				
objects measured in individual units - objects measured in groups	0	0.078	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.194	0.078	-2.507	<.05
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	-0.611	0.078	-7.879	<.001

Table O

Linear regression for behaviour data when classifier is gen

	Estimate	SE	t	*p*
Count nouns				
(Intercept)	0.25	0.008	31.843	<.001
[Cl-N]				
objects measured in individual units - objects measured in groups	0.923	0.038	24	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.038	0.038	1	.318
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.038	0	1
[Cl-A-N]				
objects measured in individual units - objects measured in groups	0.923	0.038	24	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.038	0.038	1	.318
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.038	0	1
[A-Cl-N]				
objects measured in individual units - objects measured in groups	0.923	0.038	24	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.038	0.038	1	.318
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.038	0	1
Mass nouns				
(Intercept)	0.25	0.019	13.237	<.001
[Cl-N]				
objects measured in individual units - objects measured in groups	0	0.093	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.5	0.093	-5.404	<.001
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.093	0	1
[Cl-A-N]				
objects measured in individual units - objects measured in groups	0	0.093	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.192	0.093	-2.079	<.05
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	-0.615	0.093	-6.651	<.001
[A-Cl-N]				
objects measured in individual units - objects measured in groups	0	0.093	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.385	0.093	-4.157	<.001
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	-0.231	0.093	-2.494	<.05

Table P
Linear regression for behaviour data when classifier is kuai

	Estimate	SE	t	*p*
Count nouns				
(Intercept)	0.25	0.005	46.219	<.001
[Cl-N]				
objects measured in individual units - objects measured in groups	0.952	0.027	35.392	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.024	0.027	0.885	.377
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.027	0	1
[Cl-A-N]				
objects measured in individual units - objects measured in groups	0.909	0.026	34.578	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.045	0.026	1.729	.084
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.026	0	1
[A-Cl-N]				
objects measured in individual units - objects measured in groups	0.955	0.026	36.307	<.001
objects measured in groups - entities measured in Adj-inconsistent units	0.023	0.026	0.864	.388
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.026	0	1
Mass nouns				
(Intercept)	0.25	0.014	18.11	<.001
[Cl-N]				
objects measured in individual units - objects measured in groups	0	0.067	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.5	0.067	-7.452	<.001
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	0	0.067	0	1
[Cl-A-N]				
objects measured in individual units - objects measured in groups	0	0.067	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.136	0.067	-2.032	<.05
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	-0.727	0.067	-10.839	<.001
[A-Cl-N]				
objects measured in individual units - objects measured in groups	0	0.069	0	1
objects measured in groups - entities measured in Adj-inconsistent units	-0.31	0.069	-4.507	<.001
entities measured in Adj-inconsistent units - entities measured in Adj-consistent units	-0.381	0.069	-5.547	<.001

1

2