

A vista da ponte das pessoas que vão em fila a caminho do céu: um relance sobre a composicionalidade das palavras complexas do chinês e do português

João Veloso*

jveloso@letras.up.pt

Faculdade de Letras da Universidade do Porto

Centro de Linguística da Universidade do Porto (Portugal)

RESUMO.

A composicionalidade, enquanto princípio semântico que se aplica a unidades complexas possibilitando a computação do seu significado global a partir dos significados dos seus constituintes, costuma ser aceite, pelo menos nos estudos sobre línguas indo-europeias, como uma propriedade das frases. No caso das palavras complexas, o papel da composicionalidade é admitido de forma muito limitada. Em línguas como o chinês, porém, em que a maior parte das palavras simples são monossilábicas, gerando uma elevada taxa de densidade homofónica, a construção de palavras complexas a partir de palavras monossilábicas de acordo com procedimentos de natureza composicional corresponde a um recurso muito frequente e muito sistemático. Neste trabalho, a par da necessidade de avaliar princípios como a composicionalidade de forma diferente consoante a língua em que a observamos, defenderemos que as palavras complexas do chinês – ao contrário do que sucede nas línguas indo-europeias – têm uma natureza composicional muito evidente, tal como demonstrado em inúmeros estudos anteriores.

PALAVRAS-CHAVE.

Composicionalidade; Palavra; Palavra Complexa; Chinês; Português.

ABSTRACT.

Semantic compositionality, as the principle that makes it possible to compute the global meaning of a complex unit from the specific meanings of its constituents, is usually accepted (at least when Indo-European languages are considered) as exclusive of sentences. Even in the case of complex words, the role of compositionality is accepted very limitedly only. In languages like Chinese, though, where most simple words are monosyllabic – thus, generating a high rate of homophones –, forming complex words on the basis of the compositional combination of simple words is a highly productive, systematic linguistic resource of the language. In this paper, we shall argue for the need of always assessing linguistic principles, such as compositionality, in relation to particular languages; in tandem with this tenet, we shall

* Investigação financiada pela FCT – Fundação para a Ciência e a Tecnologia através do financiamento UIDB/00022/2020 atribuído ao Centro de Linguística da Universidade do Porto.

favour the view of Chinese complex words as compositional structures (differently from what occurs in Indo-European languages, namely). This conclusion is in line with previous studies on this subject.

KEYWORDS.

Compositionality; Word; Complex Word; Chinese; Portuguese.

Dedicado à Professora Doutora Fátima Oliveira

1. Nota introdutória

Este texto representa uma primeira incursão minha na comparação descritiva entre propriedades estruturais do português e do chinês mandarim¹: o principal objeto de análise será o comportamento da composicionalidade em unidades correspondentes a *palavras* nas duas línguas contempladas.

Ainda que, a nível dos tópicos mais restritos abordados neste trabalho, esta seja a primeira vez que me aproximo de tais assuntos, este trabalho insere-se e dá seguimento a linhas de investigação sobre as quais me debrucei em anteriores ocasiões, como a definição do conceito de palavra (Veloso 2016), a circunscrição da validade de certas propostas descritivas a subconjuntos delimitados de línguas ou famílias de línguas (Veloso 2020a) e o debate entre convenção e natureza na definição das categorias gramaticais (*ibid.*). Partindo de investigações anteriores como as citadas, o presente artigo pretende cruzá-las com a minha recente condição de aprendente

1 Ao longo do texto, as designações “chinês” e “mandarim” serão usadas como equivalentes entre si para fazermos referência à norma padrão da língua oficial da China que é adotada no sistema administrativo e escolar chinês e no sistema internacional HSK de exames de proficiência em língua chinesa. Trata-se da variedade que, fortemente baseada na norma de prestígio de Pequim, é também referida, por vezes, como *Putonghua* (Duanmu 2005, 2007: 4-6, Mai *et al.* 2019: 39-40). Duanmu (2005, 2007) adota a designação de “chinês padrão” (“Standard Chinese”), que dá como equivalentes a outras, como “chinês mandarim” (“Mandarin Chinese”) ou “mandarim de Pequim” (“Beijing Mandarin”), p. ex. Todas as propostas apresentadas neste artigo têm em conta, exclusivamente, esta variedade do chinês; outras variedades da língua e outras línguas faladas na China não foram tidas em consideração no presente trabalho. Todas as transcrições do chinês apresentadas serão dadas recorrendo à romanização adotada pelo sistema Pinyin (cf. novamente, e entre outros: Duanmu 2007: 6, Mai *et al.* 2019: 39-40), de uso obrigatório nos sistemas administrativo e escolar da China Interior e de Singapura.

adulto do chinês como língua não materna. No âmbito dessa especial circunstância, posiciono-me dentro de um perfil muito particular, definido pela minha condição de linguista que, apetrechado de uma considerável porção de conhecimentos linguísticos explícitos, se coloca nos papéis simultâneos de agente, objeto e observador especialmente atento dessa experiência cognitiva, cultural e linguística muito específica. De certa forma, as atribuições de um linguista português na aprendizagem do chinês são o contexto circunstancial de fundo que alimenta o presente artigo. Com as reflexões aqui reunidas, pretendo ainda antecipar um horizonte exploratório para estudos futuros.

O texto insere-se num volume de homenagem à insigne semanticista Professora Doutora Fátima Oliveira. Não sendo eu um semanticista e nunca me tendo sequer acercado dos temas trabalhados pela homenageada e pelo seu grupo de investigação, eleger um conceito tão caro à semântica como a composicionalidade pode parecer uma ousadia e um atrevimento. Como muitas vezes disse em momentos muito diversos, nunca fui formalmente estudante da Professora Fátima Oliveira; contudo, com ela construí muitas aprendizagens fora da sala de aula e fora da relação professor(a)-aluno. Entre essas aprendizagens desenvolvidas em situação informal, guardo a recordação de um convívio e de uma conversa aparentemente trivial que ele propiciou (as pessoas que trabalham e convivem com a Professora Fátima Oliveira sabem da importância que os momentos de socialização assumem, na interação com a homenageada desta revista, para a troca e para a consolidação de ideias e de conhecimentos). Ainda a preparar o meu doutoramento, fui convidado para um jantar em casa da Fátima e do Pavel e, perante os meus desabafos sobre as agruras da solidão e da incerteza de quem escreve uma tese, a Fátima avançou com uma comparação muito interessante entre o trabalho científico e o trabalho artístico, sublinhando, num primeiro momento da conversa, a criatividade como o traço unificador de ambos. Trata-se de uma comparação que ainda hoje me acompanha e que muitas vezes me conforta, no meio de desalentos, prazos ultrapassados, angústias do papel em branco e outros tormentos parecidos. É verdade: encontrar um tema, estudá-lo a fundo e escrever sobre ele como nunca ninguém o fez antes é um trabalho de *criação* que muitas vezes requer, entre outras exigências, a mesma intuição de partida, a mesma capacidade de

imaginar para além do aparente, a mesma rebeldia, a mesma originalidade, o mesmo atrevimento e a mesma independência indispensáveis à criação artística. Como já aqui afirmei, essas palavras acompanham-me até hoje e foi com base nesse incentivo ao arrojo – arrojo científico, arrojo criativo, arrojo intelectual – que, depois de muita ponderação, resolvi, de modo independente e porventura ousado, atrever-me a escolher este tema para este artigo e a escrever sobre ele².

No desenvolvimento do tema proposto, reservei a secção 2 para algumas considerações de partida acerca do conceito de composicionalidade e da sua possível aplicação à palavra, nas duas línguas aqui consideradas. A secção 3 concentrar-se-á, numa primeira parte, nas propriedades inerentes à definição/delimitação da palavra em português e em chinês; as subsecções finais de 3 examinarão em maior detalhe as “palavras polissilábicas compostas” do chinês e a hipótese de, nesta língua, a composicionalidade assumir uma importância particular, quando comparada com o seu funcionamento em línguas como o português. Em 4, reuniremos algumas observações de síntese final e daremos uma explicação, baseada nos dados revistos ao longo de 2 e 3, para o insólito título dado a este artigo.

2. Questões centrais de partida

2.1. Composicionalidade e processamento composicional de construções linguísticas

O nosso propósito central, neste trabalho, consiste em reunir algumas observações de carácter geral sobre a possibilidade de aplicação do conceito de *composicionalidade* às unidades linguísticas que, em chinês, correspondem a palavras, comparando tais observações com o comportamento da mesma propriedade no tocante às palavras do português.

2 Foi no jantar a que aludi nestas palavras introdutórias que pela primeira vez ouvi falar, salvo erro, e pela voz da própria Professora Fátima Oliveira, do livro *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger* (Hale & Keyser (Eds.) 1993) – um volume de homenagem que, como este, festeja a vida e a carreira de alguém que se dedicou por inteiro à investigação em linguística. Foi nesse título, fazendo uma ponte entre um jantar distante no tempo e o momento presente como testemunho de muitos anos de amizade e de aprendizagem contínuas, que me inspirei para o título deste artigo.

Sabendo que o conceito de composicionalidade apresenta leituras divergentes e complexas (Spencer 1991: 88; Oliveira 2009), adotaremos aqui um entendimento sem dúvida redutor e sobressimplificado desta noção, ainda que fundamentado em definições e interpretações como as que encontramos expostas em autores como Spencer (1991: 44), Lappin (2001: 374, 388), Villalva (2008: 29, 2020: 3158-3159) ou Lieber (2016: 71, 223), entre outros. Nessas apresentações, a composicionalidade é apresentada como a propriedade de algumas unidades e construções linguísticas que torna possível a identificação de significados globais a partir da identificação e combinação dos significados dos constituintes das unidades mais complexas (cf. novamente os autores acima citados). Definições deste conceito que se aproximam desta aceção podem ser encontradas em citações como as seguintes ((1)).

- (1) Definições explícitas de “composicionalidade”: a computabilidade do significado de uma unidade complexa como o produto dos significados dos seus constituintes internos

“[...] compositionality requires that the meaning of any expression P be computable by a function which, given the meanings of P’s syntactic constituents as its arguments, yields P’s meaning as its value.”

Lappin (2001: 388)

“**Composicionalidade** é um princípio estabelecido para a semântica, segundo o qual o significado de um todo é uma função do significado das partes.”

Villalva (2008: 29)

“A **composicionalidade** é uma propriedade semântica das estruturas linguísticas, que reflete uma relação de perfeita regularidade entre o todo e as partes: a interpretação semântica do todo pode ser deduzida a partir da interpretação semântica das partes.”

Villalva (2020: 3159)

2.2. Composicionalidade e palavra

Concebida a partir das formulações referidas em 2.1 (vd. novamente as citações de (1)), a composicionalidade não se aplicaria à unidade *palavra* – sobretudo, se tivermos em mente as palavras morfologicamente simples do português –, tornando-se a sua aplicação mais óbvia em construções mais extensas e mais complexas como o sintagma e, muito especialmente, a frase.

De acordo com as referidas definições ((1)), a composicionalidade estaria necessariamente inibida no processamento de unidades morfossintáticas indivisíveis em subconstituintes dotados de significados individuais que pudessem projetar-se no significado combinado de unidades maiores. Conforme já foi mencionado, parece ser o que ocorre justamente no processamento das palavras e, mais especificamente, no processamento das palavras morfologicamente simples como as que encontramos em línguas como o português. Com efeito, para que um falante do português conheça ou reconheça os significados de “*flor*”, “*lápis*” ou “*paz*”, terá de recorrer ao seu conhecimento lexical armazenado e à informação idiossincrática associada, no seu léxico mental, a cada um desses itens: nada na sua forma (fonético-fonológica ou morfológica) fornece qualquer pista que permita aos falantes aceder diretamente ao significado dessas unidades lexicais. Tal acesso requer obrigatoriamente – e unicamente – a ativação da memória/memorização de itens dicionarizados de forma idiossincrática.

Pelo contrário, o processamento de frases consiste, fundamentalmente, em calcular (“computar” ou “deduzir”, recorrendo a termos contidos nas citações de (1)) o seu significado a partir da combinação do significado das suas partes – sintagmas, palavras, etc. Ao nível da frase, precisamente, este efeito é comprovado pela capacidade de os falantes conseguirem produzir e processar um número infinito de frases nunca antes processadas. Assim, um falante do português conseguirá atribuir significados, mais ou menos inusitados, a frases como as de (2) mesmo que nunca as tenha ouvido ou lido anteriormente, justamente por ser capaz de combinar *composicionalmente* o significado lexicalizado de cada palavra.

(2) Frases “inusitadas” passíveis de processamento composicional em português

A flor do telhado da sede da empresa onde trabalho é amarela e deita luz fluorescente todas as madrugadas entre as cinco e as seis horas da manhã.

O computador do vaidoso sobrinho da minha colega fala sozinho porque tem um programa esquisito e ultrapassado que foi instalado por engano.

2.2.1. Composicionalidade e “palavras morfologicamente complexas”

Apesar da associação, feita em 2.2, entre “significado composicional” e frase, por um lado, e “significado não composicional” e palavra, por outro, temos a registar que um número considerável de autores provenientes do campo da morfologia admite porém que a composicionalidade possa estar presente e exercer influência, pelo menos parcialmente, no processamento de um subconjunto particular de unidades identificáveis como *palavras*: as palavras morfologicamente complexas, como as derivadas afixais e os compostos morfossintáticos que podemos encontrar em português. Estas palavras, correspondendo a estruturas que combinam no seu interior unidades portadoras de significados individuais (morfemas lexicais, palavras morfológicas, p. ex.), aproximar-se-iam assim de unidades sintáticas complexas sujeitas ao referido processamento composicional. Um falante do português, assim como consegue atribuir significado às frases de (2) a partir da ativação do processamento composicional, seria capaz de interpretar semanticamente pseudoneologismos como os de (3) também a partir dessa mesma composicionalidade: conhecendo o significado (idiosincrático e memorizado no léxico mental) de cada um dos morfemas e palavras morfologicamente simples que se alinham nas construções de (3), conseguiria atribuir a estas últimas estruturas interpretações semânticas válidas na língua, tal como pretendemos aí demonstrar através das paráfrases entre parênteses retos que acompanham cada exemplo.

(3) Pseudoneologismos do português passíveis de processamento composicional

tira-solas [pessoa ou instrumento especializado em descolar solas de sapatos]

uberista [condutor de veículos Uber]

conta-vacinas [pessoa ou instrumento com a função de contar vacinas administradas]

piloto-carcereiro [pessoa com a função de pilotar um avião para transporte de prisioneiros]

caneteiro [profissional que fabrica ou conserta canetas]

mascarista [pessoa que é especialista em máscaras]

A possibilidade de processamento composicional deste tipo de palavras é admitida em diversos autores oriundos do campo da morfologia, conforme já referimos: Villalva (2008: 29, 2020: 3158 ss.) e Lieber (2016: 71), p. ex., assumem-no explicitamente. A última autora citada apresenta mesmo, a este respeito, um ponto de vista muito claro quando, na seguinte afirmação e discorrendo acerca do significado de palavra, distingue claramente entre o significado *composicional* das “palavras complexas”, predizível a partir do significado dos vários morfemas/palavras combinados no seu interior, e o significado *não composicional* (= “lexicalizado”) das “palavras simples”, relativamente às quais tal computação/predição não é possível:

“Meanings of complex words that are predictable as the sum of their parts are said to be **compositional**. Lexicalized words have meanings that are **non-compositional**.”

Lieber (2016: 71)

Contudo, mesmo os autores que defendem a possibilidade de processamento composicional de palavras morfologicamente complexas ressaltam que nem sempre a complexidade morfológica está diretamente associada à possibilidade de um imediato processamento composicional. As mesmas autoras aqui citadas registam esta reserva, que encontramos ressaltada também, p. ex., em Spencer (1991: 44, 48) e em Bochner (1993:

30). A cristalização de significados que se vão afastando da significação original; a fixação, nas expressões idiomáticas, de um significado próprio afastado do sentido original das suas palavras de base; a lexicalização progressiva de compostos etimológicos ou figurativos, etc. – nem sempre permitem fazer equivaler de forma tão direta complexidade morfológica e significado composicional. Com dois exemplos muito esclarecedores, Villalva (2020) demonstra esta inequação, ao comparar os compostos do português “*sofá-cama*” (composicionalmente processável) e “*barriga-de-freira*” (não passível de processamento composicional, porque dotado de uma lexicalização própria que confere a esta forma um significado independente e indeduzível dos significados originais das suas unidades lexicais constitutivas).

Estas reservas – que demonstram o carácter “escorregadio” de um entendimento precipitado que admitisse que todas as palavras compostas são composicionais, conforme sublinhado por Spencer (1991: 88) – parecem-nos nitidamente condensadas por Bochner (1993), quando o autor afirma que mesmo as palavras morfológicamente complexas são armazenadas no léxico mental em associação com uma porção de informação idiosincrática que não é composicionalmente computável:

“[...] the prediction that the properties of a morphologically complex word can always be predicted from the properties of its constituent morphemes, is false. [...] complex words do in fact have idiosyncratic properties, i.e., properties that cannot be predicted in any systematic way from their morphemes [...]”

Bochner (1993: 30)

* * * * *

Conforme afirmámos nas primeiras observações deste mesmo texto, o nosso objetivo central consiste principalmente em reunir alguma informação que nos permita ver, tendo presentes estas considerações iniciais relacionadas com alguma possibilidade de processamento composicional de palavras, se a composicionalidade agirá porventura do mesmo modo em duas línguas em que a construção de palavras não obedece exatamente aos

mesmos mecanismos: o português e o chinês.

Neste âmbito, pareceu-nos adequado debruçarmo-nos de seguida em algumas considerações de princípio acerca da própria noção de palavra em português e em chinês, o que será o objeto da próxima secção do texto.

3. Características e propriedades fundamentais da *palavra* em português e em chinês

Recuperamos neste momento reflexões anteriores em que defendemos que unidades linguísticas aparentemente muito estáveis – como a *palavra* – podem ser definidas, de língua para língua, com base em critérios nem sempre totalmente coincidentes entre si (Veloso 2016). Em tais trabalhos – e inspirando-nos, entre outros, nas importantes reflexões de Haspelmath (2004, 2007, 2011, 2018) acerca destas mesmas interrogações de partida e de outras questões análogas –, defendemos a necessidade de se avaliar criteriosamente até que ponto os objetos da descrição linguística são delimitados pela sua própria ontologia ou, em alternativa, por opções convencionais associadas a épocas e a escolas de pensamento linguístico muito específicas, por definição mutáveis ao longo do tempo e moderadamente arbitrárias na escolha de certas opções herdadas de tradições anteriores. Tendo em mente esta visão crítica dos catálogos das unidades linguísticas, tentaremos começar por isolar um conjunto mínimo de propriedades que possam ser aplicadas a unidades lexicais do português e do chinês e que, funcionando como fundamentalmente equivalentes entre si, possam ser aceites, nestas duas línguas, como *palavras* e, neste ponto da exposição, como “*palavras simples*”. Estes critérios correspondem, aproximadamente, aos critérios que em Veloso (2016: 48 ss.) propusemos como critérios para a identificação da “palavra mínima” das línguas flexionais.

Tais critérios mínimos são os que resumimos em (4).

- (4) Critérios mínimos para a identificação e delimitação das palavras do português e do chinês (“palavras simples”)
 - a) A palavra tem uma **forma fonético-fonológica relativamente fixa**,

- abstratamente representada no léxico teórico/mental da língua.
- b) A palavra detém **autonomia enquanto constituinte frásico**, patente em propriedades como, p. ex., a sua relativa **mobilidade**, a sua capacidade de poder corresponder ao **núcleo de um constituinte sintático** e a possibilidade de **anaforização/pronominalização**.
 - c) A palavra possui **significado ou função gramatical** estável, conforme demonstrado, p. ex., pela sua **dicionarização** convencional ou por operações de **sinonimização** ou **comutação** por formas holisticamente equivalentes.

De acordo com as tradições gramaticográficas e lexicográficas portuguesa e chinesa, estas unidades correspondem, grosso modo, às entradas dos dicionários das respetivas línguas, as quais são normalmente dotadas de definição semântica, classificação gramatical e representação gráfica convencional fixa e relativamente estável, mostrando-se passíveis de tradução, com relativa facilidade, entre ambas as línguas.

Assim, as unidades linguísticas do português e do chinês exemplificadas no Quadro 1 corresponderão às unidades que aqui tomaremos como palavras (simples) de uma língua e de outra.

QUADRO 1 – Exemplos de unidades linguísticas, equivalentes em português e em chinês, que podem ser consideradas como “palavras” em ambas as línguas

PORTUGÊS	CHINÊS
família	jiā
pessoa	rén
flor	huā
boca	kǒu
mão	shǒu
cão	gǒu
gato	māo
leite	nǎi

As propriedades que acima identificámos em (4) como os critérios mínimos para atribuímos o estatuto de palavra a uma unidade linguística, ao

nível elementar que aqui nos interessa, verificam-se em todos os exemplos do Quadro 1: todas estas unidades, de facto, possuem uma forma e um significado estabilizados nas duas línguas, todas apresentam considerável autonomia sintática e todas são passíveis de operações linguísticas básicas como a anaforização pronominal.

3.1. Algumas particularidades das palavras em chinês: palavras monossilábicas (“simples”) e palavras polissilábicas (“complexas”)

No seguimento das observações feitas na secção anterior, concentrar-nos-emos, neste momento, em algumas particularidades típicas das palavras do chinês com relevância para as questões específicas que aqui pretendemos explorar.

A primeira dessas particularidades, especialmente atestada nas *palavras simples*, é o seu carácter monossilábico. Conforme sublinhado pelas descrições gramaticais da língua, a natureza monossilábica das palavras simples do chinês é uma propriedade muito característica do seu léxico (Duanmu 2005, 2007, Mai *et al.* 2019: 71 ss.).

Dado que a combinação de segmentos consonânticos e vocálicos em monossílabos respeitando as restrições fonotáticas da língua produz, necessariamente, um número muito limitado de sequências bem formadas, o material lexical básico assim obtido, considerando única e exclusivamente as sequências lineares de consoantes e vogais, é em número bastante reduzido. Duanmu (2005) refere que os monossílabos do chinês formam um conjunto relativamente escasso de 400 itens (considerando somente, conforme já referido, os alinhamentos segmentais). Para poder, com base neste material de partida relativamente exíguo, construir um léxico capaz de exprimir todos os significados veiculáveis pelos enunciados linguísticos, evitado um número demasiado elevado de formas homófonas, o chinês amplia este repertório fundamental através de alguns procedimentos básicos, de que aqui destacamos:

1) atribuição de papel distintivo a propriedades autosegmentais como o tom: vd. exemplos como yú ‘peixe’ / yù ‘jade’, wáng ‘rei’ / wǎng ‘rede’, xiě ‘escrever’ / xiē ‘alguns’ ou diàn ‘elétrico/eletricidade’ / diǎn ‘ponto/em ponto’;

2) desambiguação da palavra monossilábica pela anteposição de um “classificador” que restringe a classe semântica da palavra;

3) criação de palavras polissilábicas que resultam sempre da combinação de palavras monossilábicas “primitivas” (aspeto que discutiremos em pormenor na secção seguinte);

Como dissemos, o foco da nossa atenção incidirá, neste trabalho, no recurso referido em 3), ou seja, na criação de palavras complexas polissilábicas a partir de monossílabos lexicais.

Segundo o já citado estudo de Duanmu (2005), mesmo que associado os 400 monossílabos do chinês aos contrastes tonais admitidos pela fonologia da língua, o número de unidades lexicais simples e monossilábicas assim obtido não ultrapassa os 1.300 itens, o que, tendo em vista somente as palavras mais correntes do chinês, deixa ainda uma margem substancial para a existência de um elevado número de formas homófonas, cujo processamento (sem contar, p. ex., com a ativação do conhecimento da escrita) pode acarretar riscos de ambiguidade que, como atrás referido, só o contexto pode frequentemente ajudar a resolver. Na sua introdução à fonologia do chinês, o mesmo autor refere a “densidade de homófonos” – isto é, a associação de nuvens de múltiplos significados a uma mesma palavra monossilábica –, estimando que cada monossílabo segmental do chinês (e não considerando sequer os vocabulários técnicos, eruditos, literários e poéticos, nem as aceções altamente especializadas, figurativas ou metafóricas de muitas palavras monossilábicas) possa corresponder a pelo menos 15 significados (= palavras) diferentes:

“According to a text corpus of over 45 million Chinese character tokens [...], there are more than 6,000 different Chinese characters, most of which are monosyllabic words. This means that each SC [=Standard Chinese] syllable represents about 15 words excluding tones, or five words including tones.”

Duanmu (2007: 93)

Esta característica do acervo monossilábico do chinês motiva, assim, e conforme já referido, a necessidade de combinar esses mesmos monossílabos em unidades polissilábicas³. Trata-se, como vimos, de um recurso utilizado pelo chinês para ampliar o seu léxico por forma a poder exprimir todos os significados possíveis e necessários a partir de um repertório limitado de entradas lexicais monossilábicas. A formação deste tipo de palavras é um processo sistemático e produtivo e é a partir de tal regularidade que a tradição gramatical chinesa distingue explícita e categoricamente entre palavras monossilábicas “simples” e palavras polissilábicas “complexas”, sendo estas últimas construídas principalmente a partir do alinhamento, dentro de uma mesma unidade lexical, dos monossílabos fundamentais do léxico (Mai et al. 2019: 71 ss.).

3.1.1. As palavras polissilábicas do chinês

Trataremos, nesta secção, das palavras polissilábicas do chinês, procurando ver se e como é possível aceitar o papel especial da composicionalidade, nesta língua, como propriedade relevante para a construção e o processamento de tais unidades, cujo estatuto de palavra obedece às propriedades assumidas em (4).

Para tanto, principiaremos por reunir alguns exemplos de palavras polissilábicas do chinês formadas por concatenação de palavras monossilábicas – as unidades que Mai et al. (2019: 75) categorizam no grupo das “palavras compostas por associação de morfemas”⁴ (“PCAM”), designação que seria mais adequadamente substituída por “palavras compostas por associação de palavras simples”, dado que os constituintes destas palavras reúnem os requisitos de palavra enunciados em (4). Tais exemplos são reunidos no Quadro 2, no qual apresentamos numa coluna à parte (terceira coluna), para cada palavra polissilábica, as palavras monossilábicas a partir das quais ela se constrói.

3 “No Chinês Clássico, a esmagadora maioria das palavras é monossilábica. Por isso, a homofonia é muito frequente. No Chinês Moderno, usam-se principalmente palavras polissilábicas, nomeadamente dissilábicas, o que reduz consideravelmente o número de homófonos. [...]” (Mai et al. 2019: 74).

4 Outros processos de construção de palavras “complexas” que visam ampliar o inventário lexical do chinês, segundo Mai et al. (2019: 74 ss.), são, designadamente, a repetição silábica, a associação coordenativa e a associação determinativa, de que aqui não nos ocuparemos.

QUADRO 2 – Exemplos de palavras polissilábicas compostas em chinês por combinação de palavras monossilábicas simples (“PCAM” – “Palavras Compostas por Associação de Morfemas”, *ap. Mai et al. 2019*)

PCAM	Significado lexicalizado da PCAM	Monossilabos combinados na PCAM
shǒujī	‘telemóvel’	shǒu ‘mão’ jī ‘máquina’
fēiji	avião	fēi ‘voar’ jī ‘máquina’
diànnǎo	computador	diàn ‘elétrico/eletricidade’ nǎo ‘cérebro’
diànyǐng	filme	diàn ‘elétrico/eletricidade’ yǐng ‘sombra’
diànshì	televisão	diàn ‘elétrico/eletricidade’ shì ‘imagem’
tiānqì	clima	tiān ‘céu’ qì ‘vapor/névoa’
míngtiān	amanhã	míng ‘brilho’ tiān ‘céu’
dàxué	universidade	dà ‘grande’ xué ‘estudo/escola’
huǒchē	comboio	huǒ ‘fogo’ chē ‘carro/viatura/meio de transporte’
shūdiàn	livraria	shū ‘livro’ diàn ‘loja’

Observando atentamente a formação das palavras compostas exemplificadas no Quadro 2, verificamos que em todas estas, como já foi referido, são alinhadas palavras que podem corresponder a palavras simples: trata-se das palavras monossilábicas apresentadas na 3ª coluna do Quadro 2.

A possibilidade de processamento composicional das PCAM parece-nos minimamente demonstrada ao vermos que o resultado de cada uma delas é, pelo menos parcialmente, recomponível (“computável” ou “deduzível”, recuperando aqui os termos de Lappin (2001) e Villalva (2020) encontrados nas citações em (1)) a partir da combinação dos significados das palavras monossilábicas encontradas no interior de cada uma destas

PCAM. Paráfrases como as que apresentamos na 4ª coluna do Quadro 3, explicitando e reconstruindo a relação parte-todo em que estas palavras polissilábicas parecem assentar, ilustram a possibilidade de concebermos o significado destas unidades, pelo menos em parte, como um significado *composicional*.

QUADRO 3. A composicionalidade das PCAM em chinês: o reconhecimento dos significados de cada palavra monossilábica e a sua projeção composicional no significado da palavra composta

PCAM	Significado lexicalizado da PCAM	Monossilabos combinados na PCAM	Paráfrase interpretativa da composicionalidade da palavra composta
shǒujī	'telemóvel'	shǒu 'mão' jī 'máquina'	[máquina de usar na mão]
fēijī	avião	fēi 'voar' jī 'máquina'	[máquina de voar/ máquina voadora]
diànnǎo	computador	diàn 'elétrico/eletricidade' nǎo 'cérebro'	[cérebro elétrico]
diànyǐng	filme	diàn 'elétrico/eletricidade' yǐng 'sombra'	[sombra projetada pela eletricidade]
diànshì	televisão	diàn 'elétrico/eletricidade' shì 'imagem'	[imagem fabricada/projetada pela eletricidade]
tiānqì	clima	tiān 'céu' qì 'vapor/névoa'	[o vapor/a névoa do céu]
míngtiān	amanhã	míng 'brilho' tiān 'céu'	[o brilho que virá [novamente] do céu]
dàxué	universidade	dà 'grande' xué 'estudo/escola'	[escola para estudos superiores [= grandes]]
huǒchē	comboio	huǒ 'fogo' chē 'carro/viatura/meio de transporte'	[viatura movida a fogo]
shūdiàn	livraria	shū 'livro' diàn 'loja'	[loja/estabelecimento de livros]

Estes exemplos, assim explicados, mostram-nos que, em chinês, o significado da palavra composta estabelece sempre algum tipo de relação com o significado de cada uma das palavras simples nela incluídas. Tendo em atenção que este recurso é fundamental, em chinês, para ampliar o

léxico – evitando o problema da densidade de homófonos (Duanmu 2005) – e que, nesta língua, ele corresponde a um processo muito recorrente e sistemático, responsável pela geração de um número muito elevado de palavras correntes, devemos refletir melhor, relativamente a esta língua, sobre a exclusão da composicionalidade do domínio da palavra. Ou seja: parecendo certo que, em línguas como o português (com sistemas fonológicos e lexicais diferentes dos do chinês), a composicionalidade, nos termos percorridos principalmente na secção 2, dificilmente pode ser associada à palavra – mesmo à palavra complexa –, em línguas como o chinês, de acordo com os dados revistos até ao momento, a composicionalidade parece adquirir uma importância e, sobretudo, uma produtividade muito próprias na formação e no processamento das palavras complexas. A importância da composicionalidade ao nível da palavra complexa do chinês, de resto, constitui um tópico reconhecido e amplamente investigado em inúmeros estudos de carácter experimental, como, a título de exemplo, Arcodia (2007), Yang *et al.* (2012), Lazaridou *et al.* (2013), Su & Kim (2014), Zhu & Fellbaum (2015) e Zhan *et al.* (2020).

Também as gramáticas pedagógicas do chinês, evidenciando a importância desta propriedade das palavras nesta língua, adotam frequentemente a convenção de apresentarem, para cada palavra, um “sentido literal” (o significado individualizado de cada uma das palavras simples constitutivas) e o seu “sentido global” (o significado *composicional* da palavra composta, como tal registado em dicionários e seguido, p. ex., na tradução para outras línguas). Este é o procedimento utilizado, p. ex., pela *Gramática de Língua Chinesa para Falantes de Português* de Mai *et al.* (2019).

4. Breves considerações finais

Perante esta análise sumária dos dados lexicais aqui apresentados e relacionando-a com as questões centrais do estudo enunciadas nas secções iniciais, parece-nos lícito concluir que, até certo ponto pelo menos, a composicionalidade – que, em línguas como o português, parece importante sobretudo para a construção e o processamento do sintagma e da frase –

desempenha um papel importante para a construção e o processamento de unidades lexicais do chinês coincidentes com o conceito de palavra complexa. Como tal, algumas das reservas encontradas junto de autores como Spencer (1991) ou Bochner (1993), alertando para uma impossibilidade de se estender de forma automática o conceito de complexidade morfológica da palavra ao conceito de composicionalidade, farão talvez mais sentido em línguas da família indo-europeia do que em línguas com as características tipológicas do chinês. Se tivermos em conta as características fonológicas particulares do léxico básico monossilábico do chinês referidas sobretudo ao longo da secção 3 (muito particularmente, o elevado número de homófonos monossilábicos e a densidade de homófonos referida por Duanmu 2005) e se lembrarmos que alguns potenciais problemas emergentes dessas limitações são resolvidos através de processos muito sistemáticos de combinação composicional de palavras monossilábicas (em número diminuto) em palavras polissilábicas (em número potencialmente infinito), teremos de aceitar que a relevância da composicionalidade na construção de palavras em chinês, por responder a objetivos específicos da língua no quadro de um sistema lexical-gramatical muito particular, deve ser merecedora de uma conceção radicalmente diferente daquela de que é objeto quando pensamos em línguas como o português e a generalidade das línguas indo-europeias. No chinês, as evidências do carácter composicional das palavras complexas identificam, pelos motivos aduzidos, um processo de lexicalização com uma importância diferente daquela que tem em português e noutras línguas aparentadas com o português. Num estudo mais alargado, esta verificação poderia ainda estender-se ao processamento da palavra escrita, incluindo o de palavras monossilábicas simples também.

Combinada com a compreensão dos sentidos metafóricos e figurativos de que esta língua faz também um uso muito extensivo, a composicionalidade lexical do chinês leva-nos a compreender melhor os mecanismos que levam à construção de compostos como a palavra que, literalmente traduzida, se encontraria na tradução chinesa no título deste trabalho: para designar as passagens aéreas que permitem aos peões atravessar as estradas de maior movimento através de plataformas elevadas sobre as vias de circulação das cidades, o chinês forjou a palavra *rénxíngtiānqiáo*. No seu “sentido literal”, a palavra, traduzível composicionalmente por “passagem aérea para peões”,

significaria algo como “a ponte por onde as pessoas seguem em fila em direção ao céu”:

<i>rén</i>	<i>xíng</i>	<i>tiān</i>	<i>qiáo</i>
‘pessoa’	‘ir em fila’	‘céu’	‘ponte’

Estas pontes pedonais, muito comuns nas cidades chinesas, permitem-nos muitas vezes um relance sobre as dinâmicas urbanas e populacionais dessas cidades. Neste artigo, onde pretendemos estabelecer pontes de vários tipos (entre disciplinas, entre línguas, entre tradições gramaticais, entre conceitos, processos e unidades parcialmente iguais e parcialmente diferentes), propusemo-nos também levar a cabo um primeiro relance sobre questões que nos suscitam, neste momento, algum interesse sobre as dinâmicas do chinês na construção do seu léxico.

REFERÊNCIAS

- Arcodia, G. F. (2007). Chinese: A Language of Compound Words? In F. Montermini, G. Boyé, & N. Hathout (Eds.), *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse* (pp. 79-90). Cascadilla Proceedings.
- Bochner, H. (1993). *Simplicity in Generative Morphology*. De Gruyter.
- Coulmas, F. (2003). *Writing Systems. An introduction to their linguistic analysis*. Cambridge University Press.
- Daniels, P. (2010). Writing in the World and Linguistics. *Papers of the Berkeley Linguistics Society* 36, 61-90.
- Duanmu, S. (2005). Chinese (Mandarin), Phonology of. In K. Brown, (Eds.), *Encyclopedia of Language and Linguistics* (2ª ed.). Elsevier.
- Duanmu, S. (2007). *The Phonology of Standard Chinese* (2ª ed.). Oxford University Press.
- Hale, K. L., & Keyser, S. J. (Eds.). (1993). *The View From Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. The MIT Press.
- Haspelmath, M. (2004). Does linguistic explanation presuppose linguistic description? *Studies in Language*, 28(3), 554-579.
- Haspelmath, M. (2007). Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology*, 11(1), 119-132.
- Haspelmath, M. (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica*, 45(1), 31-80.
- Haspelmath, M. (2018). How comparative concepts and descriptive linguistic categories are different. In D. Van Olmen, T. Mortelmans, & F. Brisard (Eds.), *Aspects of linguistic variation* (pp. 83-114). De Gruyter Mouton.
- Lappin, S. (2001). An Introduction to Formal Semantics. In M. Aronoff, & J. Rees-Miller (Eds.), *The Handbook of Linguistics* (pp. 369-393). Blackwell.
- Lazaridou, A., Marelli, M., Zamparelli, R., & Baroni, M. (2013). Compositionally Derived Representations of Morphologically Complex Words in Distributional Semantics. In H. Schuetze, P. Fung, & M. Poesio (Eds.), *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (pp. 1517-1526). Association for Computational Linguistics.
- Li, W., Gaffney, J. S., & Packard, J. L. (Eds.). (2002). *Chinese Children's Reading Acquisition. Theoretical and Pedagogical Issues*. Kluwer.
- Lieber, R. (2016). *Introducing Morphology* (2ª ed.). Cambridge University Press.
- Liu, F., Lu, H., Lo, C., & Neubig, G. (2017). Learning Character-level Compositionality

- with Visual Features. In R. Barzilay, & M.-Y. Kan (Eds.), *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (pp. 2059-2068). Association for Computational Linguistics.
- Mai, R., Morais, C., & Pereira, U. (2019). *Gramática de Língua Chinesa para Falantes de Português*. Instituto Confúcio, Universidade de Aveiro.
- Miller, K. F. (2002). Children's Early Understanding of Writing and Language: The Impact of Characters and Alphabetic Orthographies. In: W. Li, J. S. Gaffney, & J. Packard (Eds), *Chinese Children's Reading Acquisition. Theoretical and Pedagogical Issues* (pp. 17-29). Kluwer.
- Nagy, W. E., Kuo-Kealoha, A., Xinchun, W., Wenling, L., Anderson, R. C., & Xi, C. (2002). The Role of Morphological Awareness in Learning to Read Chinese. In: W. Li, J. S. Gaffney, & J. Packard (Eds), *Chinese Children's Reading Acquisition. Theoretical and Pedagogical Issues* (pp. 59-86). Kluwer.
- Oliveira, F. (2009). Alguns Caminhos da Semântica. In M. A. D. C. Coutinho, & A. Fiéis (Orgs.), *Textos Seleccionados. XXIV Encontro Nacional da Associação Portuguesa de Linguística* (pp. 23-41). Associação Portuguesa de Linguística.
- Spencer, A. (1991). *Morphological Theory. An Introduction to Word Structure in Generative Grammar*. Blackwell.
- Su, X., & Kim, Y.-S. (2014). Semantic radical knowledge and word recognition in Chinese for Chinese as foreign language learners. *Reading in a Foreign Language*, 26(1), 131-152.
- Veloso, J. (2005). A língua na escrita e a escrita da língua. Algumas considerações gerais sobre transparência e opacidade fonémicas na escrita do português e outras questões. *Da Investigação às Práticas. Estudos de Natureza Educacional*, VI(1), 49-69.
- Veloso, J. (2007). *Da Influência do Conhecimento Ortográfico sobre o Conhecimento Fonológico. Estudo Longitudinal de um Grupo de Crianças Falantes Nativas do Português Europeu*. Lincom Europa
- Veloso, J. (2010). Primeiras produções escritas e operações metafonológicas explícitas como pistas para a caracterização inferencial do conhecimento fonológico. *Cadernos de Educação*, (35), 19-50.
- Veloso, J. (2016). Verba Manent. *A palavra como unidade pertinente para a descrição linguística do português e de outras línguas flexionais* (Vol. 2, pp. 160). EDIPUCRS.
- Veloso, J. (2019). Phonology and Writing: Can we look at written productions to "see the unseeable" in phonology? *Loquens*, 6(1), 1-12.
- Veloso, J. (2020a, outubro 28-30). *Consoantes e Vogais? Primitivamente Iguais?*

- [Conferência plenária]. XXXVI Encontro Nacional da Associação Portuguesa de Linguística.
- Veloso, J. (2020b). Conhecimento ortográfico e representações fonológicas em português. In S. N. Salomão (Coord.), *Temas da Língua Portuguesa: do Pluricentrismo à Didática* (pp. 91-103). Nuova Cultura.
- Veloso, J. (2022, no prelo). *Fonologia e Ortografia do Português Europeu*. EDULOG 2022.
- Villalva, A. (2008). *Morfologia do Português*. Universidade Aberta.
- Villalva, A. (2020). Composição. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, & A. Mendes (Orgs.), *Gramática do Português* (Vol. 3, pp. 3151,3210). Fundação Calouste Gulbenkian.
- Yang, W., Dai, W., & Gao, L. (2012). A Study on Chinese Continuous Four-character Collocations and Their Translation into English Strategies. *Theory and Practice in Language Studies*, 2(5), 995-1002.
- Zhan, W., Wang, J., Chen, L., & Huang, H. (2020). Form and Meaning Representation of Chinese Constructions. *Fundamental Issues on Constructicography*. *Sinica Venetiana*, 6, 305-338.
- Zhu, F., & Fellbaum, C. (2015). Quantifying Fixedness and Compositionality in Chinese Idioms. *International Journal of Lexicography*, 28(3), 338–350.